Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2442092 A1 2002/10/17

(21) 2 442 092

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2002/03/26
(87) Date publication PCT/PCT Publication Date: 2002/10/17
(85) Entrée phase nationale/National Entry: 2003/09/25
(86) N° demande PCT/PCT Application No.: US 2002/009187
(87) N° publication PCT/PCT Publication No.: 2002/081494
(30) Priorités/Priorities: 2001/03/26 (09/817,879) US; 2001/06/08 (60/296,876) US; 2001/06/08 (09/877,478) US; 2001/10/24 (60/335,059) US; 2001/12/05 (60/337,055) US

(51) Cl.Int.[7]/Int.Cl.[7] C07H 21/02, A01K 67/033, A01K 67/00, A01K 67/027, C07H 21/04, G01N 33/00

(71) Demandeur/Applicant:
RIBOZYME PHARMACEUTICALS, INC., US

(72) Inventeurs/Inventors:
BLATT, LAWRENCE, US;
MACEJAK, DENNIS, US;
MCSWIGGEN, JAMES, US;
MORRISSEY, DAVID, US;
PAVCO, PAMELA, US;
LEE, PATRICE, US;
...

(74) Agent: SMART & BIGGAR

(54) Titre : INHIBITION REGULEE PAR DES OLIGONUCLEOTIDES DE LA REPLICATION DU VIRUS DE L'HEPATITE B ET DU VIRUS DE L'HEPATITE C

(54) Title: OLIGONUCLEOTIDE MEDIATED INHIBITION OF HEPATITIS B VIRUS AND HEPATITIS C VIRUS REPLICATION

(57) Abrégé/Abstract:
The present invention relates to nucleic acid molecules, including antisense and enzymatic nucleic acid molecules, such as hammerhead ribozymes, DNAzymes, Inozymes, Zinzymes, Amberzymes, and G-cleaver ribozymes, which modulate the synthesis, expression and/or stability of an HCV or HBV RNA and methods for their use alone or in combination with other therapies. In addition, nucleic acid decoy molecules and aptamers that bind to HBV reverse transcriptase and/or HBV reverse transcriptase primer sequences and methods for their use alone or in combination with other therapies, are disclosed. Oligonucleotides that specifically bind the Enhancer I region of HBV DNA are further disclosed. The present invention further relates to the use of nucleic acids, such as decoy and aptamer molecules of the invention, to modulate the expression of Hepatitis B virus (HBV) genes and HBV viral replication. Furthermore, HBV animal models and methods of use are disclosed, including methods of screening for compounds and/or potential therapies directed against HBV. The present invention also relates to compounds, including enzymatic nucleic acid molecules, ribozymes, DNAzymes, nuclease activating compounds and chimeras such as 2',5'-adenylates, that modulate the expression and/or replication of hepatitis C virus (HCV).

(72) **Inventeurs(suite)/Inventors(continued):** DRAPER, KENNETH, US; ROBERTS, ELISABETH, US

CA 02442092 2003-09-25

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 October 2002 (17.10.2002)

PCT

(10) International Publication Number
WO 02/081494 A1

(51) International Patent Classification[7]: C07H 21/02, 21/04, A01K 67/00, 67/033, 67/027, G01N 33/00

(21) International Application Number: PCT/US02/09187

(22) International Filing Date: 26 March 2002 (26.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/817,879	26 March 2001 (26.03.2001)	US
09/877,478	8 June 2001 (08.06.2001)	US
60/296,876	8 June 2001 (08.06.2001)	US
60/335,059	24 October 2001 (24.10.2001)	US
60/337,055	5 December 2001 (05.12.2001)	US

(71) Applicant *(for all designated States except US)*: **RIBOZYME PHARMACEUTICALS, INC.** [US/US]; 2950 Wilderness Place, Boulder, CO 80301 (US).

(71) Applicants and
(72) Inventors: **BLATT, Lawrence** [US/US]; 2176 Riverside Lane, Boulder, CO 80304 (US). **MACEJAK, Dennis** [US/US]; 6595 Union Street, Arvada, CO 80004 (US). **MCSWIGGEN, James** [US/US]; 4866 Franklin Dr., Boulder, CO 80301 (US). **MORRISSEY, David** [US/US]; 4769 Tanglewood Trail, Boulder, CO 80301 (US). **PAVCO, Pamela** [US/US]; 705 Barberry Circle, LaFayette, CO 80026 (US). **LEE, Patrice** [US/US]; 91 Autumn Court, Erie, CO 80516 (US). **DRAPER, Kenneth** [US/US]; 4791 Cougar Creek Trail, Reno, NV 89509 (US). **ROBERTS, Elisabeth** [US/US]; 10025 Eliot Circle, Federal Heights, CO 80260 (US).

(74) Agent: **TERPSTRA, Anita, J.**; McDonnell Boehnen Hulbert & Berghoff, 32nd Floor, 300 South Wacker Drive, Chicago, IL 60606 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: OLIGONUCLEOTIDE MEDIATED INHIBITION OF HEPATITIS B VIRUS AND HEPATITIS C VIRUS REPLICATION

(57) Abstract: The present invention relates to nucleic acid molecules, including antisense and enzymatic nucleic acid molecules, such as hammerhead ribozymes, DNAzymes, Inozymes, Zinzymes, Amberzymes, and G-cleaver ribozymes, which modulate the synthesis, expression and/or stability of an HCV or HBV RNA and methods for their use alone or in combination with other therapies. In addition, nucleic acid decoy molecules and aptamers that bind to HBV reverse transcriptase and/or HBV reverse transcriptase primer sequences and methods for their use alone or in combination with other therapies, are disclosed. Oligonucleotides that specifically bind the Enhancer I region of HBV DNA are further disclosed. The present invention further relates to the use of nucleic acids, such as decoy and aptamer molecules of the invention, to modulate the expression of Hepatitis B virus (HBV) genes and HBV viral replication. Furthermore, HBV animal models and methods of use are disclosed, including methods of screening for compounds and/or potential therapies directed against HBV. The present invention also relates to compounds, including enzymatic nucleic acid molecules, ribozymes, DNAzymes, nuclease activating compounds and chimeras such as 2',5'-adenylates, that modulate the expression and/or replication of hepatitis C virus (HCV).

WO 02/081494 A1

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME	1	DE	2
CONTENANT LES PAGES	1	À	193

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME	1	OF	2
CONTAINING PAGES	1	TO	193

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHIER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

DESCRIPTION

OLIGONUCLEOTIDE MEDIATED INHIBITION OF HEPATITIS B VIRUS AND HEPATITIS C VIRUS REPLICATION

Background Of The Invention

This patent application claims priority from Blatt et al., USSN (09/817,879), filed March 26, 2001, which is a continuation-in-part of Blatt et al., USSN (09/740,332), filed December 18, 2000, which is a continuation-in-part of Blatt et al., USSN (09/611,931), filed July 7, 2000, which is a continuation-in-part of Blatt et al., 09/504,321, filed February 15, 2000, which is a continuation-in-part of Blatt et al., USSN 09/274,553, filed March 23, 1999, which is a continuation-in-part of Blatt et al., USSN 09/257,608, filed February 24, 1999 (abandoned), which claims priority from Blatt et al., USSN 60/100,842, filed September 18, 1998, and McSwiggen et al., USSN 60/083,217 filed April 27, 1998; all of these earlier applications are entitled "ENZYMATIC NUCLEIC ACID TREATMENT OF DISEASES OR CONDITIONS RELATED TO HEPATITIS C VIRUS INFECTION". This patent application also claims priority from Draper et al., USSN 09/877,478 filed June 8, 2001, which is a continuation-in-part of Draper et al., USSN (09/696,347), filed October 24, 2000, which is a continuation-in-part of Draper et al., USSN (09/636,385), filed August 9, 2000, which is a continuation in part of Draper et al., USSN (09/531,025), filed March 20, 2000, which is a continuation in part of Draper, USSN (09/436,430), filed November 8, 1999, which is a continuation of USSN (08/193,627), filed February 7, 1994, now US patent No. 6,017,756, which is a continuation of USSN (07/882,712), filed May 14, 1992, now abandoned; all of these earlier applications are entitled "METHOD AND REAGENT FOR INHIBITING HEPATITIS B VIRUS REPLICATION". This patent application also claims priority from Macejak et al., USSN (60/335,059), filed October 24, 2001, Macejak et al., USSN (60/296,876), filed June 8, 2001, and Morrissey et al., USSN (60/337,055), filed December 5, 2001. These applications are hereby incorporated by reference herein in their entireties, including the drawings.

The present invention concerns compounds, compositions, and methods for the study, diagnosis, and treatment of degenerative and disease states related to hepatitis B virus (HBV) and hepatitis C virus (HCV) infection, replication and gene expression. Specifically, the invention relates to nucleic acid molecules used to modulate expression of HBV and HCV. In

addition, the instant invention relates to methods, models and systems for screening inhibitors of HBV and HCV replication and propagation.

The following is a discussion of relevant art pertaining to hepatitis B virus (HBV) and hepatitis C virus (HCV). The discussion is not meant to be complete and is provided only for understanding of the invention that follows. The summary is not an admission that any of the work described below is prior art to the claimed invention.

In 1989, the Hepatitis C Virus (HCV) was determined to be an RNA virus and was identified as the causative agent of most non-A non-B viral Hepatitis (Choo *et al., Science.* 1989; 244:359-362). Unlike retroviruses such as HIV, HCV does not go though a DNA replication phase and no integrated forms of the viral genome into the host chromosome have been detected (Houghton *et al., Hepatology* 1991;14:381-388). Rather, replication of the coding (plus) strand is mediated by the production of a replicative (minus) strand leading to the generation of several copies of plus strand HCV RNA. The genome consists of a single, large, open-reading frame that is translated into a polyprotein (Kato *et al., FEBS Letters.* 1991; 280: 325-328). This polyprotein subsequently undergoes post-translational cleavage, producing several viral proteins (Leinbach *et al., Virology.* 1994: 204:163-169).

Examination of the 9.5-kilobase genome of HCV has demonstrated that the viral nucleic acid can mutate at a high rate (Smith *et al.,Mol. Evol.* 1997 45:238-246). This rate of mutation has led to the evolution of several distinct genotypes of HCV that share approximately 70% sequence identity (Simmonds *et al., J. Gen. Virol.* 1994;75 :1053-1061). It is important to note that these sequences are evolutionarily quite distant. For example, the genetic identity between humans and primates such as the chimpanzee is approximately 98%. In addition, it has been demonstrated that an HCV infection in an individual patient is composed of several distinct and evolving quasispecies that have 98% identity at the RNA level. Thus, the HCV genome is hypervariable and continuously changing. Although the HCV genome is hypervariable, there are 3 regions of the genome that are highly conserved. These conserved sequences occur in the 5' and 3' non-coding regions as well as the 5'-end of the core protein coding region and are thought to be vital for HCV RNA replication as well as translation of the HCV polyprotein. Thus, therapeutic agents that target these conserved HCV genomic regions can have a significant impact over a wide range of HCV genotypes. Moreover, it is unlikely that drug resistance will occur with enzymatic nucleic acids specific to conserved regions of the HCV genome. In contrast, therapeutic modalities that target inhibition of enzymes such as the viral proteases or helicase are likely to result in the selection for drug resistant strains since the RNA for these viral encoded enzymes is located in the hypervariable portion of the HCV genome.

After initial exposure to HCV, the patient experiences a transient rise in liver enzymes, which indicates the occurrence of inflammatory processes (Alter *et al.*, IN: Seeff LB, Lewis JH, eds. *Current Perspectives in Hepatology.* New York: Plenum Medical Book Co; 1989:83-89). This elevation in liver enzymes will occur at least 4 weeks after the initial exposure and can last for up to two months (Farci *et al., New England Journal of Medicine.* 1991:325:98-104). Prior to the rise in liver enzymes, it is possible to detect HCV RNA in the patient's serum using RT-PCR analysis (Takahashi *et al., American Journal of Gastroenterology.* 1993:88:2:240-243). This stage of the disease is called the acute stage and usually goes undetected since 75% of patients with acute viral hepatitis from HCV infection are asymptomatic. The remaining 25% of these patients develop jaundice or other symptoms of hepatitis.

Acute HCV infection is a benign disease, however, and as many as 80% of acute HCV patients progress to chronic liver disease as evidenced by persistent elevation of serum alanine aminotransferase (ALT) levels and by continual presence of circulating HCV RNA (Sherlock, *Lancet* 1992; 339:802). The natural progression of chronic HCV infection over a 10 to 20 year period leads to cirrhosis in 20 to 50% of patients (Davis *et al., Infectious Agents and Disease* 1993;2:150:154) and progression of HCV infection to hepatocellular carcinoma has been well documented (Liang *et al., Hepatology.* 1993; 18:1326-1333; Tong *et al., Western Journal of Medicine*, 1994; Vol. 160, No. 2: 133-138). There have been no studies that have determined sub-populations that are most likely to progress to cirrhosis and/or hepatocellular carcinoma, thus all patients have equal risk of progression.

It is important to note that the survival for patients diagnosed with hepatocellular carcinoma is only 0.9 to 12.8 months from initial diagnosis (Takahashi *et al., American Journal of Gastroenterology.* 1993:88:2:240-243). Treatment of hepatocellular carcinoma with chemotherapeutic agents has not proven effective and only 10% of patients will benefit from surgery due to extensive tumor invasion of the liver (Trinchet *et al., Presse Medicine.* 1994:23:831-833). Given the aggressive nature of primary hepatocellular carcinoma, the only viable treatment alternative to surgery is liver transplantation (Pichlmayr *et al., Hepatology.* 1994:20:33S-40S).

Upon progression to cirrhosis, patients with chronic HCV infection present with clinical features, which are common to clinical cirrhosis regardless of the initial cause (D'Amico *et al., Digestive Diseases and Sciences.* 1986;31:5: 468-475). These clinical features can include: bleeding esophageal varices, ascites, jaundice, and encephalopathy (Zakim D, Boyer TD. Hepatology a textbook of liver disease. Second Edition Volume 1. 1990 W.B. Saunders Company. Philadelphia). In the early stages of cirrhosis, patients are classified as compensated, meaning that although liver tissue damage has occurred, the patient's liver is still able to detoxify metabolites in the blood-stream. In addition, most

patients with compensated liver disease are asymptomatic and the minority with symptoms report only minor symptoms such as dyspepsia and weakness. In the later stages of cirrhosis, patients are classified as decompensated meaning that their ability to detoxify metabolites in the bloodstream is diminished and it is at this stage that the clinical features described above will present.

In 1986, D'Amico *et al.* described the clinical manifestations and survival rates in 1155 patients with both alcoholic and viral associated cirrhosis (D'Amico *supra*). Of the 1155 patients, 435 (37%) had compensated disease although 70% were asymptomatic at the beginning of the study. The remaining 720 patients (63%) had decompensated liver disease with 78% presenting with a history of ascites, 31% with jaundice, 17% had bleeding and 16% had encephalopathy. Hepatocellular carcinoma was observed in six (.5%) patients with compensated disease and in 30 (2.6%) patients with decompensated disease.

Over the course of six years, the patients with compensated cirrhosis developed clinical features of decompensated disease at a rate of 10% per year. In most cases, ascites was the first presentation of decompensation. In addition, hepatocellular carcinoma developed in 59 patients who initially presented with compensated disease by the end of the six-year study.

With respect to survival, the D'Amico study indicated that the five-year survival rate for all patients on the study was only 40%. The six-year survival rate for the patients who initially had compensated cirrhosis was 54%, while the six-year survival rate for patients who initially presented with decompensated disease was only 21%. There were no significant differences in the survival rates between the patients who had alcoholic cirrhosis and the patients with viral related cirrhosis. The major causes of death for the patients in the D'Amico study were liver failure in 49%; hepatocellular carcinoma in 22%; and, bleeding in 13% (D'Amico *supra*).

Chronic Hepatitis C is a slowly progressing inflammatory disease of the liver, mediated by a virus (HCV) that can lead to cirrhosis, liver failure and/or hepatocellular carcinoma over a period of 10 to 20 years. In the US, it is estimated that infection with HCV accounts for 50,000 new cases of acute hepatitis in the United States each year (NIH Consensus Development Conference Statement on Management of Hepatitis C March 1997). The prevalence of HCV in the United States is estimated at 1.8% and the CDC places the number of chronically infected Americans at approximately 4.5 million people. The CDC also estimates that up to 10,000 deaths per year are caused by chronic HCV infection. The prevalence of HCV in the United States is estimated at 1.8% and the CDC places the number of chronically infected Americans at approximately 4.5 million people. The CDC also estimates that up to 10,000 deaths per year are caused by chronic HCV infection.

Numerous well controlled clinical trials using interferon (IFN-alpha) in the treatment of chronic HCV infection have demonstrated that treatment three times a week results in lowering of serum ALT values in approximately 50% (range 40% to 70%) of patients by the end of 6 months of therapy (Davis *et al., New England Journal of Medicine* 1989; 321:1501-1506; Marcellin *et al., Hepatology.* 1991; 13:393-397; Tong *et al., Hepatology* 1997:26:747-754; Tong *et al., Hepatology* 1997 26(6): 1640-1645). However, following cessation of interferon treatment, approximately 50% of the responding patients relapsed, resulting in a "durable" response rate as assessed by normalization of serum ALT concentrations of approximately 20 to 25%.

In recent years, direct measurement of the HCV RNA has become possible through use of either the branched-DNA or Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) analysis. In general, the RT-PCR methodology is more sensitive and leads to more accurate assessment of the clinical course (Tong *et al., supra*). Studies that have examined six months of type 1 interferon therapy using changes in HCV RNA values as a clinical endpoint have demonstrated that up to 35% of patients will have a loss of HCV RNA by the end of therapy (Marcellin *et al., supra*). However, as with the ALT endpoint, about 50% of the patients relapse six months following cessation of therapy resulting in a durable virologic response of only 12% (Marcellin *et al., supra*). Studies that have examined 48 weeks of therapy have demonstrated that the sustained virological response is up to 25% (NIH consensus statement: 1997). Thus, standard of care for treatment of chronic HCV infection with type 1 interferon is now 48 weeks of therapy using changes in HCV RNA concentrations as the primary assessment of efficacy (Hoofnagle *et al., New England Journal of Medicine* 1997; 336(5) 347-356).

Side effects resulting from treatment with type 1 interferons can be divided into four general categories, which include 1. Influenza-like symptoms; 2. Neuropsychiatric; 3. Laboratory abnormalities; and, 4. Miscellaneous (Dusheiko *et al.,* Journal of Viral Hepatitis. 1994:1:3-5). Examples of influenza-like symptoms include; fatigue, fever; myalgia; malaise; appetite loss; tachycardia; rigors; headache and arthralgias. The influenza-like symptoms are usually short-lived and tend to abate after the first four weeks of dosing (Dushieko *et al., supra*). Neuropsychiatric side effects include: irritability, apathy; mood changes; insomnia; cognitive changes and depression. The most important of these neuropsychiatric side effects is depression and patients who have a history of depression should not be given type 1 interferon. Laboratory abnormalities include; reduction in myeloid cells including granulocytes, platelets and to a lesser extent red blood cells. These changes in blood cell counts rarely lead to any significant clinical sequellae (Dushieko *et al., supra*). In addition, increases in triglyceride concentrations and elevations in serum alanine and aspartate aminotransferase concentration have been observed. Finally, thyroid abnormalities have been reported. These thyroid abnormalities are usually reversible after cessation of interferon

therapy and can be controlled with appropriate medication while on therapy. Miscellaneous side effects include nausea; diarrhea; abdominal and back pain; pruritus; alopecia; and rhinorrhea. In general, most side effects will abate after 4 to 8 weeks of therapy (Dushieko *et al., supra*).

Type 1 Interferon is a key constituent of many treatment programs for chronic HCV infection. Treatment with type 1 interferon induces a number of genes and results in an antiviral state within the cell. One of the genes induced is 2', 5' oligoadenylate synthetase, an enzyme that synthesizes short 2', 5' oligoadenylate (2-5A) molecules. Nascent 2-5A subsequently activates a latent RNase, RNase L, which in turn nonspecifically degrades viral RNA.

Chronic hepatitis B is caused by an enveloped virus, commonly known as the hepatitis B virus or HBV. HBV is transmitted via infected blood or other body fluids, especially saliva and semen, during delivery, sexual activity, or sharing of needles contaminated by infected blood. Individuals may be "carriers" and transmit the infection to others without ever having experienced symptoms of the disease. Persons at highest risk are those with multiple sex partners, those with a history of sexually transmitted diseases, parenteral drug users, infants born to infected mothers, "close" contacts or sexual partners of infected persons, and healthcare personnel or other service employees who have contact with blood. Transmission is also possible via tattooing, ear or body piercing, and acupuncture; the virus is also stable on razors, toothbrushes, baby bottles, eating utensils, and some hospital equipment such as respirators, scopes and instruments. There is no evidence that HBsAg positive food handlers pose a health risk in an occupational setting, nor should they be excluded from work. Hepatitis B has never been documented as being a food-borne disease. The average incubation period is 60 to 90 days, with a range of 45 to 180; the number of days appears to be related to the amount of virus to which the person was exposed. However, determining the length of incubation is difficult, since onset of symptoms is insidious. Approximately 50% of patients develop symptoms of acute hepatitis that last from 1 to 4 weeks. Two percent or less of these individuals develop fulminant hepatitis resulting in liver failure and death.

The determinants of severity include: (1) The size of the dose to which the person was exposed; (2) the person's age with younger patients experiencing a milder form of the disease; (3) the status of the immune system with those who are immunosuppressed experiencing milder cases; and (4) the presence or absence of co-infection with the Delta virus (hepatitis D), with more severe cases resulting from co-infection. In symptomatic cases, clinical signs include loss of appetite, nausea, vomiting, abdominal pain in the right upper quadrant, arthralgia, and tiredness/loss of energy. Jaundice is not experienced in all

cases, however, jaundice is more likely to occur if the infection is due to transfusion or percutaneous serum transfer, and it is accompanied by mild pruritus in some patients. Bilirubin elevations are demonstrated in dark urine and clay-colored stools, and liver enlargement may occur accompanied by right upper-quadrant pain. The acute phase of the disease may be accompanied by severe depression, meningitis, Guillain-Barré syndrome, myelitis, encephalitis, agranulocytosis, and/or thrombocytopenia.

Hepatitis B is generally self-limiting and will resolve in approximately 6 months. Asymptomatic cases can be detected by serologic testing, since the presence of the virus leads to production of large amounts of HBsAg in the blood. This antigen is the first and most useful diagnostic marker for active infections. However, if HBsAg remains positive for 20 weeks or longer, the person is likely to remain positive indefinitely and is now a carrier. While only 10% of persons over age 6 who contract HBV become carriers, 90% of infants infected during the first year of life do so.

Hepatitis B virus (HBV) infects over 300 million people worldwide (Imperial, 1999, *Gastroenterol. Hepatol.*, 14 (suppl), S1-5). In the United States, approximately 1.25 million individuals are chronic carriers of HBV as evidenced by the fact that they have measurable hepatitis B virus surface antigen HBsAg in their blood. The risk of becoming a chronic HBsAg carrier is dependent upon the mode of acquisition of infection as well as the age of the individual at the time of infection. For those individuals with high levels of viral replication, chronic active hepatitis with progression to cirrhosis, liver failure and hepatocellular carcinoma (HCC) is common, and liver transplantation is the only treatment option for patients with end-stage liver disease from HBV.

The natural progression of chronic HBV infection over a 10 to 20 year period leads to cirrhosis in 20-to-50% of patients and progression of HBV infection to hepatocellular carcinoma has been well documented. There have been no studies that have determined sub-populations that are most likely to progress to cirrhosis and/or hepatocellular carcinoma, thus all patients have equal risk of progression.

It is important to note that the survival for patients diagnosed with hepatocellular carcinoma is only 0.9 to 12.8 months from initial diagnosis (Takahashi *et al.*, 1993, *American Journal of Gastroenterology*, 88, 240-243). Treatment of hepatocellular carcinoma with chemotherapeutic agents has not proven effective and only 10% of patients will benefit from surgery due to extensive tumor invasion of the liver (Trinchet *et al.*, 1994,*Presse Medicine*, 23, 831-833). Given the aggressive nature of primary hepatocellular carcinoma, the only viable treatment alternative to surgery is liver transplantation (Pichlmayr *et al.*, 1994, *Hepatology.*, 20, 33S-40S).

Upon progression to cirrhosis, patients with chronic HCV and HBV infection present with clinical features, which are common to clinical cirrhosis regardless of the initial cause (D'Amico *et al.*, 1986, *Digestive Diseases and Sciences*, 31, 468-475). These clinical features may include: bleeding esophageal varices, ascites, jaundice, and encephalopathy (Zakim D, Boyer TD. *Hepatology a textbook of liver disease*, Second Edition Volume 1. 1990 W.B. Saunders Company. Philadelphia). In the early stages of cirrhosis, patients are classified as compensated, meaning that although liver tissue damage has occurred, the patient's liver is still able to detoxify metabolites in the blood-stream. In addition, most patients with compensated liver disease are asymptomatic and the minority with symptoms report only minor symptoms such as dyspepsia and weakness. In the later stages of cirrhosis, patients are classified as decompensated meaning that their ability to detoxify metabolites in the bloodstream is diminished and it is at this stage that the clinical features described above will present.

In 1986, D'Amico *et al.* described the clinical manifestations and survival rates in 1155 patients with both alcoholic and viral associated cirrhosis (D'Amico *supra*). Of the 1155 patients, 435 (37%) had compensated disease although 70% were asymptomatic at the beginning of the study. The remaining 720 patients (63%) had decompensated liver disease with 78% presenting with a history of ascites, 31% with jaundice, 17% had bleeding and 16% had encephalopathy. Hepatocellular carcinoma was observed in six (0.5%) patients with compensated disease and in 30 (2.6%) patients with decompensated disease.

Over the course of six years, the patients with compensated cirrhosis developed clinical features of decompensated disease at a rate of 10% per year. In most cases, ascites was the first presentation of decompensation. In addition, hepatocellular carcinoma developed in 59 patients who initially presented with compensated disease by the end of the six-year study.

With respect to survival, the D'Amico study indicated that the five-year survival rate for all patients on the study was only 40%. The six-year survival rate for the patients who initially had compensated cirrhosis was 54% while the six-year survival rate for patients who initially presented with decompensated disease was only 21%. There were no significant differences in the survival rates between the patients who had alcoholic cirrhosis and the patients with viral related cirrhosis. The major causes of death for the patients in the D'Amico study were liver failure in 49%; hepatocellular carcinoma in 22%; and, bleeding in 13% (D'Amico *supra*).

Hepatitis B virus is a double-stranded circular DNA virus. It is a member of the Hepadnaviridae family. The virus consists of a central core that contains a core antigen (HBcAg) surrounded by an envelope containing a surface protein/surface antigen (HBsAg)

and is 42 nm in diameter. It also contains an e antigen (HBeAg), which, along with HBcAg and HBsAg, is helpful in identifying this disease.

In HBV virions, the genome is found in an incomplete double-stranded form. HBV uses a reverse transcriptase to transcribe a positive-sense full length RNA version of its genome back into DNA. This reverse transcriptase also contains DNA polymerase activity and thus begins replicating the newly synthesized minus-sense DNA strand. However, it appears that the core protein encapsidates the reverse-transcriptase/polymerase before it completes replication.

From the free-floating form, the virus must first attach itself specifically to a host cell membrane. Viral attachment is one of the crucial steps that determines host and tissue specificity. However, currently there are no *in vitro* cell-lines that can be infected by HBV. There are some cells lines, such as HepG2, which can support viral replication only upon transient or stable transfection using HBV DNA.

After attachment, fusion of the viral envelope and host membrane must occur to allow the viral core proteins containing the genome and polymerase to enter the cell. Once inside, the genome is translocated to the nucleus where it is repaired and cyclized.

The complete closed circular DNA genome of HBV remains in the nucleus and gives rise to four transcripts. These transcripts initiate at unique sites but share the same 3'-ends. The 3.5-kb pregenomic RNA serves as a template for reverse transcription and also encodes the nucleocapsid protein and polymerase. A subclass of this transcript with a 5'-end extension codes for the precore protein that, after processing, is secreted as HBV e antigen. The 2.4-kb RNA encompasses the pre-S1 open reading frame (ORF) that encodes the large surface protein. The 2.1-kb RNA encompasses the pre-S2 and S ORFs that encode the middle and small surface proteins, respectively. The smallest transcript (~0.8-kb) codes for the X protein, a transcriptional activator.

Multiplication of the HBV genome begins within the nucleus of an infected cell. RNA polymerase II transcribes the circular HBV DNA into greater-than-full length mRNA. Since the mRNA is longer than the actual complete circular DNA, redundant ends are formed. Once produced, the pregenomic RNA exits the nucleus and enters the cytoplasm.

The packaging of pregenomic RNA into core particles is triggered by the binding of the HBV polymerase to the 5' epsilon stem-loop. RNA encapsidation is believed to occur as soon as binding occurs. The HBV polymerase also appears to require associated core protein in order to function. The HBV polymerase initiates reverse transcription from the 5' epsilon stem-loop three to four base pairs at which point the polymerase and attached nascent DNA

are transferred to the 3' copy of the DR1 region. Once there, the (-)DNA is extended by the HBV polymerase while the RNA template is degraded by the HBV polymerase RNAse H activity. When the HBV polymerase reaches the 5' end, a small stretch of RNA is left undigested by the RNAse H activity. This segment of RNA is comprised of a small sequence just upstream and including the DR1 region. The RNA oligomer is then translocated and annealed to the DR2 region at the 5' end of the (-)DNA. It is used as a primer for the (+)DNA synthesis which is also generated by the HBV polymerase. It appears that the reverse transcription as well as plus strand synthesis may occur in the completed core particle.

Since the pregenomic RNA is required as a template for DNA synthesis, this RNA is an excellent target for nucleic acid based therapeutics. Nucleoside analogues that have been documented to modulate HBV replication target the reverse transcriptase activity needed to convert the pregenomic RNA into DNA. Nucleic acid decoy and aptamer modulation of HBV reverse transcriptase would be expected to result in a similar modulation of HBV replication.

Current therapeutic goals of treatment are three-fold: to eliminate infectivity and transmission of HBV to others, to arrest the progression of liver disease and improve the clinical prognosis, and to prevent the development of hepatocellular carcinoma (HCC).

Interferon alpha use is the most common therapy for HBV; however, recently Lamivudine (3TC®) has been approved by the FDA. Interferon alpha (IFN-alpha) is one treatment for chronic hepatitis B. The standard duration of IFN-alpha therapy is 16 weeks, however, the optimal treatment length is still poorly defined. A complete response (HBV DNA negative HBeAg negative) occurs in approximately 25% of patients. Several factors have been identified that predict a favorable response to therapy including:, High ALT, low HBV DNA, being female, and heterosexual orientation.

There is also a risk of reactivation of the hepatitis B virus even after a successful response, this occurs in around 5% of responders and normally occurs within 1 year.

Side effects resulting from treatment with type 1 interferons can be divided into four general categories including: Influenza-like symptoms, neuropsychiatric, laboratory abnormalities, and other miscellaneous side effects. Examples of influenza-like symptoms include, fatigue, fever, myalgia, malaise, appetite loss, tachycardia, rigors, headache and arthralgias. The influenza-like symptoms are usually short-lived and tend to abate after the first four weeks of dosing (Dusheiko *et al.*, 1994, *Journal of Viral Hepatitis*, 1, 3-5). Neuropsychiatric side effects include irritability, apathy, mood changes, insomnia, cognitive

changes, and depression. Laboratory abnormalities include the reduction of myeloid cells, including granulocytes, platelets and to a lesser extent, red blood cells. These changes in blood cell counts rarely lead to any significant clinical sequellae. In addition, increases in triglyceride concentrations and elevations in serum alanine and aspartate aminotransferase concentration have been observed. Finally, thyroid abnormalities have been reported. These thyroid abnormalities are usually reversible after cessation of interferon therapy and can be controlled with appropriate medication while on therapy. Miscellaneous side effects include nausea, diarrhea, abdominal and back pain, pruritus, alopecia, and rhinorrhea. In general, most side effects will abate after 4 to 8 weeks of therapy (Dushieko *et al., supra*).

Lamivudine (3TC®) is a nucleoside analogue, which is a very potent and specific inhibitor of HBV DNA synthesis. Lamivudine has recently been approved for the treatment of chronic Hepatitis B. Unlike treatment with interferon, treatment with 3TC® does not eliminate the HBV from the patient. Rather, viral replication is controlled and chronic administration results in improvements in liver histology in over 50% of patients. Phase III studies with 3TC®, showed that treatment for one year was associated with reduced liver inflammation and a delay in scarring of the liver. In addition, patients treated with Lamivudine (100mg per day) had a 98 percent reduction in hepatitis B DNA and a significantly higher rate of seroconversion, suggesting disease improvements after completion of therapy. However, stopping of therapy resulted in a reactivation of HBV replication in most patients. In addition recent reports have documented 3TC® resistance in approximately 30% of patients.

Current therapies for treating HBV infection, including interferon and nucleoside analogues, are only partially effective. In addition, drug resistance to nucleoside analogues is now emerging, making treatment of chronic Hepatitis B more difficult. Thus, a need exists for effective treatment of this disease that utilizes antiviral modulators that work by mechanisms other than those currently utilized in the treatment of both acute and chronic hepatitis B infections.

Welch *et al., Gene Therapy* 1996 3(11): 994-1001 describe *in vitro* an *in vivo* studies with two vector expressed hairpin ribozymes targeted against hepatitis C virus.

Sakamoto *et al., J. Clinical Investigation* 1996 98(12): 2720-2728 describe intracellular cleavage of hepatitis C virus RNA and inhibition of viral protein translation by certain vector expressed hammerhead ribozymes.

Lieber *et al., J. Virology* 1996 70(12): 8782-8791 describe elimination of hepatitis C virus RNA in infected human hepatocytes by adenovirus-mediated expression of certain hammerhead ribozymes.

Ohkawa *et al.*, 1997, *J. Hepatology*, 27; 78-84, describe *in vitro* cleavage of HCV RNA and inhibition of viral protein translation using certain *in vitro* transcribed hammerhead ribozymes.

Barber *et al.*, International PCT Publication No. *WO 97/32018*, describe the use of an adenovirus vector to express certain anti-hepatitis C virus hairpin ribozymes.

Kay *et al.*, International PCT Publication No. *WO 96/18419*, describe certain recombinant adenovirus vectors to express anti-HCV hammerhead ribozymes.

Yamada *et al.*, Japanese Patent Application No. *JP 07231784* describe a specific poly-(L)-lysine conjugated hammerhead ribozyme targeted against HCV.

Draper, U.S. Patent Nos. 5,610,054 and 5,869,253, describes enzymatic nucleic acid molecules capable of inhibiting replication of HCV.

Macejak. *et al.*, 2000, *Hepatology*, 31, 769-776, describe enzymatic nucleic acid molecules capable of inhibiting replication of HCV.

Weifeng and Torrence, 1997, *Nucleosides and Nucleotides*, 16, 7-9, describe the synthesis of 2-5A antisense chimeras with various non-nucleoside components.

Torrence *et al.*, US patent No. 5,583,032 describe targeted cleavage of RNA using an antisense oligonulceotide linked to a 2',5'-oligoadenylate activator of RNase L.

Suhadolnik and Pfleiderer, US patent Nos. 5,863,905; 5,700,785; 5,643,889; 5,556,840; 5,550,111; 5,405,939; 5,188,897; 4,924,624; and 4,859,768 describe specific internucleotide phosphorothioate 2',5'-oligoadenlyates and 2',5'-oligoadenlyate conjugates.

Budowsky *et al.*, US patent No. 5,962,431 describe a method of treating papillomavirus using specific 2',5'-oligoadenylates.

Torrence *et al.*, International PCT publication No. WO 00/14219, describe specific peptide nucleic acid 2',5'-oligoadenylate chimeric molecules.

Stinchcomb *et al.*, US patent No. 5,817,796, describe C-myb ribozymes having 2'-5'-Linked Adenylate Residues.

Draper, US patent No. 6,017,756, describes the use of ribozymes for the inhibition of Hepatitis B Virus.

Passman *et al.*, 2000, *Biochem. Biophys. Res. Commun.*, 268(3), 728-733.; Gan *et al.*, 1998, *J. Med. Coll. PLA*, 13(3), 157-159.; Li *et al.*, 1999, *Jiefangjun Yixue Zazhi*, 24(2), 99-

101.; Putlitz *et al.*, 1999, *J. Virol.*, 73(7), 5381-5387.; Kim *et al.*, 1999, *Biochem. Biophys. Res. Commun.*, 257(3), 759-765.; Xu *et al.*, 1998, *Bingdu Xuebao*, 14(4), 365-369.; Welch *et al.*, 1997, *Gene Ther.*, 4(7), 736-743.; Goldenberg *et al.*, 1997, International PCT publication No. WO 97/08309, Wands *et al.*, 1997, *J. of Gastroenterology and Hepatology*, 12(suppl.), S354-S369.; Ruiz *et al.*, 1997, *BioTechniques*, 22(2), 338-345.; Gan *et al.*, 1996, *J. Med. Coll. PLA*, 11(3), 171-175.; Beck and Nassal, 1995, *Nucleic Acids Res.*, 23(24), 4954-62.; Goldenberg, 1995, International PCT publication No. WO 95/22600.; Xu *et al.*, 1993, *Bingdu Xuebao*, 9(4), 331-6.; Wang *et al.*, 1993, *Bingdu Xuebao*, 9(3), 278-80, all describe ribozymes that are targeted to cleave a specific HBV target site.

Hunt *et al.*, US patent No. 5,859,226, describes specific non-naturally occurring oligonucleotide decoys intended to inhibit the expression of MHC-II genes through binding of the RF-X transcription factor, that can inhibit the expression of certain HBV and CMV viral proteins.

Kao *et al.*, International PCT Publication No. WO 00/04141, describes linear single stranded nucleic acid molecules capable of specifically binding to viral polymerases and inhibiting the activity of the viral polymerase.

Lu, International PCT Publication No. WO 99/20641, describes specific triplex-forming oligonucleotides used in treating HBV infection.

SUMMARY OF THE INVENTION

This invention relates to enzymatic nucleic acid molecules that can disrupt the function of RNA species of hepatitis B virus (HBV), hepatitis C virus (HCV) and/or those RNA species encoded by HBV or HCV. In particular, applicant provides enzymatic nucleic acid molecules capable of specifically cleaving HBV RNA or HCV RNA and describes the selection and function thereof. Such enzymatic nucleic acid molecules can be used to treat diseases and disorders associated with HBV and HCV infection.

In one embodiment, the invention features an enzymatic nucleic acid molecule that specifically cleaves RNA derived from hepatitis B virus (HBV), wherein the enzymatic nucleic acid molecule comprises sequence defined as Seq. ID No. 10887.

In another embodiment, the invention features a composition comprising an enzymatic nucleic acid molecule of the invention and a pharmaceutically acceptable carrier.

In another embodiment, the invention features a mammalian cell, for example a human cell, comprising an enzymatic nucleic acid molecule contemplated by the invention.

In one embodiment, the invention features a method for the treatment of cirrhosis, liver failure or hepatocellular carcinoma comprising administering to a patient an enzymatic nucleic acid molecule of the invention under conditions suitable for the treatment.

In another embodiment, the invention features a method for the treatment of a patient having a condition associated with HBV and/or HCV infection, comprising contacting cells of said patient with an enzymatic nucleic acid molecule of the invention.

In another embodiment, the invention features a method for the treatment of a patient having a condition associated with HBV and/or HCV infection, comprising contacting cells of said patient with an enzymatic nucleic acid molecule of the invention and further comprising the use of one or more drug therapies, for example, type I interferon or 3TC® (lamivudine), under conditions suitable for said treatment. In another embodiment, the other therapy is administered simultaneously with or separately from the enzymatic nucleic acid molecule.

In another embodiment, the invention features a method for inhibiting HBV and/or HCV replication in a mammalian cell comprising administering to the cell an enzymatic nucleic acid molecule of the invention under conditions suitable for the inhibition.

In yet another embodiment, the invention features a method of cleaving a separate HBV and/or HCV RNA comprising contacting an enzymatic nucleic acid molecule of the invention with the separate RNA under conditions suitable for the cleavage of the separate RNA.

In one embodiment, cleavage by an enzymatic nucleic acid molecule of the invention is carried out in the presence of a divalent cation, for example Mg2+.

In another embodiment, the enzymatic nucleic acid molecule of the invention is chemically synthesized.

In another embodiment, the type I interferon contemplated by the invention is interferon alpha, interferon beta, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, polyethylene glycol consensus interferon.

In one embodiment, the invention features a composition comprising type I interferon and an enzymatic nucleic acid molecule of the invention and a pharmaceutically acceptable carrier.

In another embodiment, the invention features a method of administering to a cell, for example a mammalian cell or human cell, an enzymatic nucleic acid molecule of the

invention independently or in conjunction with other therapeutic compounds, such as type I interferon or 3TC® (lamivudine), comprising contacting the cell with the enzymatic nucleic acid molecule under conditions suitable for the administration.

In another embodiment, administration of an enzymatic nucleic acid molecule of the invention is in the presence of a delivery reagent, for example a lipid, cationic lipid, phospholipid, or liposome.

In another embodiment, the invention features novel nucleic acid-based techniques such as enzymatic nucleic acid molecules and antisense molecules and methods for their use to down regulate or inhibit the expression of HBV RNA and/or replication of HBV.

In another embodiment, the invention features novel nucleic acid-based techniques such as enzymatic nucleic acid molecules and antisense molecules and methods for their use to down regulate or inhibit the expression of HCV RNA and/or replication of HCV.

In one embodiment, the invention features the use of one or more of the enzymatic nucleic acid-based techniques to down-regulate or inhibit the expression of the genes encoding HBV and/or HCV viral proteins. Specifically, the invention features the use of enzymatic nucleic acid-based techniques to specifically down-regulate or inhibit the expression of the HBV and/or HCV viral genome.

In another embodiment, the invention features nucleic acid-based inhibitors (*e.g.*, enzymatic nucleic acid molecules (ribozymes), antisense nucleic acids, triplex DNA, decoys, siRNA, aptamers, and antisense nucleic acids containing RNA cleaving chemical groups) and methods for their use to down regulate or inhibit the expression of RNA (*e.g.*, HBV and/or HCV) capable of progression and/or maintenance of hepatitis, hepatocellular carcinoma, cirrhosis, and/or liver failure.

In one embodiment, nucleic acid molecules of the invention are used to treat HBV infected cells or an HBV infected patient wherein the HBV is resistant or the patient does not respond to treatment with 3TC® (Lamivudine), either alone or in combination with other therapies under conditions suitable for the treatment.

In yet another embodiment, the invention features the use of an enzymatic nucleic acid molecule, preferably in the hammerhead, NCH (Inozyme), G-cleaver, amberzyme, zinzyme, and/or DNAzyme motif, to inhibit the expression of HBV and/or HCV RNA.

The enzymatic nucleic acid molecules described herein exhibit a high degree of specificity for only the viral mRNA in infected cells. Nucleic acid molecules of the instant invention targeted to highly conserved sequence regions allow the treatment of many strains

of human HBV and/or HCV with a single compound. No treatment presently exists which specifically attacks expression of the viral gene(s) that are responsible for transformation of hepatocytes by HBV and/or HCV.

The enzymatic nucleic acid-based modulators of HBV and HCV expression are useful for the prevention of the diseases and conditions including HBV and HCV infection, hepatitis, cancer, cirrhosis, liver failure, and any other diseases or conditions that are related to the levels of HBV and/or HCV in a cell or tissue.

Preferred target sites are genes required for viral replication, a non-limiting example includes genes for protein synthesis, such as the 5' most 1500 nucleotides of the HBV pregenomic mRNAs. For sequence references, see Renbao *et al.*, 1987, *Sci. Sin.*, 30, 507. This region controls the translational expression of the core protein (C), X protein (X) and DNA polymerase (P) genes and plays a role in the replication of the viral DNA by serving as a template for reverse transcriptase. Disruption of this region in the RNA results in deficient protein synthesis as well as incomplete DNA synthesis (and inhibition of transcription from the defective genomes). Targeting sequences 5' of the encapsidation site can result in the inclusion of the disrupted 3' RNA within the core virion structure and targeting sequences 3' of the encapsidation site can result in the reduction in protein expression from both the 3' and 5' fragments.

Alternative regions outside of the 5' most 1500 nucleotides of the pregenomic mRNA also make suitable targets for enzymatic nucleic acid mediated inhibition of HBV replication. Such targets include the mRNA regions that encode the viral S gene. Selection of particular target regions will depend upon the secondary structure of the pregenomic mRNA. Targets in the minor mRNAs can also be used, especially when folding or accessibility assays in these other RNAs reveal additional target sequences that are unavailable in the pregenomic mRNA species.

A desirable target in the pregenomic RNA is a proposed bipartite stem-loop structure in the 3'-end of the pregenomic RNA which is believed to be critical for viral replication (Kidd and Kidd-Ljunggren, 1996. *Nuc. Acid Res*. 24:3295-3302). The 5'end of the HBV pregenomic RNA carries a *cis*-acting encapsidation signal, which has inverted repeat sequences that are thought to form a bipartite stem-loop structure. Due to a terminal redundancy in the pregenomic RNA, the putative stem-loop also occurs at the 3'-end. While it is the 5' copy which functions in polymerase binding and encapsidation, reverse transcription actually begins from the 3' stem-loop. To start reverse transcription, a 4 nt primer which is covalently attached to the polymerase is made, using a bulge in the 5' encapsidation signal as template. This primer is then shifted, by an unknown mechanism, to the DR1 primer binding site in the 3' stem-loop structure, and reverse transcription proceeds

from that point. The 3' stem-loop, and especially the DR1 primer binding site, appear to be highly effective targets for ribozyme intervention.

Sequences of the pregenomic RNA are shared by the mRNAs for surface, core, polymerase, and X proteins. Due to the overlapping nature of the HBV transcripts, all share a common 3'-end. Enzymatic nucleic acids targeting of this common 3'-end will thus cleave the pregenomic RNA as well as all of the mRNAs for surface, core, polymerase and X proteins.

At least seven basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds in *trans* (and thus can cleave other RNA molecules) under physiological conditions. **Table I** summarizes some of the characteristics of these enzymatic RNA molecules. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets. Thus, a single enzymatic nucleic acid molecule is able to cleave many molecules of target RNA. In addition, the enzymatic nucleic acid is a highly specific inhibitor of gene expression, with the specificity of inhibition depending not only on the base-pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a an enzymatic nucleic acid molecule.

The enzymatic nucleic acid molecules that cleave the specified sites in HBV-specific RNAs represent a novel therapeutic approach to treat a variety of pathologic indications, including, HBV infection, hepatitis, hepatocellular carcinoma, tumorigenesis, cirrhosis, liver failure and other conditions related to the level of HBV.

In one of the preferred embodiments of the inventions described herein, the enzymatic nucleic acid molecule is formed in a hammerhead or hairpin motif, but can also be formed in the motif of a hepatitis delta virus, group I intron, group II intron or RNase P RNA (in association with an RNA guide sequence), *Neurospora* VS RNA, DNAzymes, NCH cleaving motifs, or G-cleavers. Examples of such hammerhead motifs are described by Dreyfus, *supra,* Rossi *et al.*, 1992, *AIDS Research and Human Retroviruses* 8, 183. Examples of hairpin motifs are described by Hampel *et al.*, EP0360257, Hampel and Tritz, 1989

Biochemistry 28, 4929, Feldstein *et al.*, 1989, *Gene* 82, 53, Haseloff and Gerlach, 1989, *Gene*, 82, 43, Hampel *et al.*, 1990 *Nucleic Acids Res.* 18, 299; and Chowrira & McSwiggen, US. Patent No. 5,631,359. The hepatitis delta virus motif is described by Perrotta and Been, 1992 *Biochemistry* 31, 16. The RNase P motif is described by Guerrier-Takada *et al.*, 1983 *Cell* 35, 849; Forster and Altman, 1990, *Science* 249, 783; and Li and Altman, 1996, *Nucleic Acids Res.* 24, 835. The *Neurospora* VS RNA ribozyme motif is described by Collins (Saville and Collins, 1990 *Cell* 61, 685-696; Saville and Collins, 1991 *Proc. Natl. Acad. Sci. USA* 88, 8826-8830; Collins and Olive, 1993 *Biochemistry* 32, 2795-2799; and Guo and Collins, 1995, *EMBO. J.* 14, 363). Group II introns are described by Griffin *et al.*, 1995, *Chem. Biol.* 2, 761; Michels and Pyle, 1995, *Biochemistry* 34, 2965; and Pyle *et al.*, International PCT Publication No. WO 96/22689. The Group I intron is described by Cech *et al.*, U.S. Patent 4,987,071. DNAzymes are described by Usman *et al.*, International PCT Publication No. WO 95/11304; Chartrand *et al.*, 1995, *NAR* 23, 4092; Breaker *et al.*, 1995, *Chem. Bio.* 2, 655; and Santoro *et al.*, 1997, *PNAS* 94, 4262. NCH cleaving motifs are described in Ludwig & Sproat, International PCT Publication No. WO 98/58058; and G-cleavers are described in Kore *et al.*, 1998, *Nucleic Acids Research* 26, 4116-4120 and Eckstein *et al.*, International PCT Publication No. WO 99/16871. Additional motifs include the Aptazyme (Breaker *et al.*, WO 98/43993), Amberzyme (Class I motif; **Figure 3**; Beigelman *et al.*, International PCT publication No. WO 99/55857) and Zinzyme (Beigelman *et al.*, International PCT publication No. WO 99/55857), all these references are incorporated by reference herein in their totalities, including drawings and can also be used in the present invention. These specific motifs are not limiting in the invention and those skilled in the art will recognize that all that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule (Cech *et al.*, U.S. Patent No. 4,987,071).

In preferred embodiments of the present invention, a nucleic acid molecule, *e.g.*, an antisense molecule, a triplex DNA, or a ribozyme, is 13 to 100 nucleotides in length, *e.g.*, in specific embodiments 35, 36, 37, or 38 nucleotides in length (*e.g.*, for particular ribozymes or antisense). In particular embodiments, the nucleic acid molecule is 15-100, 17-100, 20-100, 21-100, 23-100, 25-100, 27-100, 30-100, 32-100, 35-100, 40-100, 50-100, 60-100, 70-100, or 80-100 nucleotides in length. Instead of 100 nucleotides being the upper limit on the length ranges specified above, the upper limit of the length range can be, for example, 30, 40, 50, 60, 70, or 80 nucleotides. Thus, for any of the length ranges, the length range for particular embodiments has lower limit as specified, with an upper limit as specified which is greater than the lower limit. For example, in a particular embodiment, the length range can be 35-50 nucleotides in length. All such ranges are expressly included. Also in particular

embodiments, a nucleic acid molecule can have a length which is any of the lengths specified above, for example, 21 nucleotides in length.

Exemplary enzymatic nucleic acid molecules of the invention targeting HBV are shown in **Tables V-XI**. For example, enzymatic nucleic acid molecules of the invention are preferably between 15 and 50 nucleotides in length, more preferably between 25 and 40 nucleotides in length, *e.g.*, 34, 36, or 38 nucleotides in length (for example see Jarvis *et al.*, 1996, *J. Biol. Chem.*, 271, 29107-29112). Exemplary DNAzymes of the invention are preferably between 15 and 40 nucleotides in length, more preferably between 25 and 35 nucleotides in length, *e.g.*, 29, 30, 31, or 32 nucleotides in length (see for example Santoro *et al.*, 1998, *Biochemistry*, 37, 13330-13342; Chartrand *et al.*, 1995, Nucleic Acids Research, 23, 4092-4096). Exemplary antisense molecules of the invention are preferably between 15 and 75 nucleotides in length, more preferably between 20 and 35 nucleotides in length, *e.g.*, 25, 26, 27, or 28 nucleotides in length (see for example Woolf *et al.*, 1992, *PNAS.*, 89, 7305-7309; Milner *et al.*, 1997, *Nature Biotechnology*, 15, 537-541). Exemplary triplex forming oligonucleotide molecules of the invention are preferably between 10 and 40 nucleotides in length, more preferably between 12 and 25 nucleotides in length, *e.g.*, 18, 19, 20, or 21 nucleotides in length (see for example Maher *et al.*, 1990, *Biochemistry*, 29, 8820-8826; Strobel and Dervan, 1990, *Science*, 249, 73-75). Those skilled in the art will recognize that all that is required is for the nucleic acid molecule are of length and conformation sufficient and suitable for the nucleic acid molecule to catalyze a reaction contemplated herein. The length of the nucleic acid molecules of the instant invention are not limiting within the general limits stated.

In a preferred embodiment, the invention provides a method for producing a class of nucleic acid–based gene inhibiting agents which exhibit a high degree of specificity for the RNA of a desired target. For example, the enzymatic nucleic acid molecule is preferably targeted to a highly conserved sequence region of target RNAs encoding HBV proteins (specifically HBV RNA) such that specific treatment of a disease or condition can be provided with either one or several nucleic acid molecules of the invention. Such nucleic acid molecules can be delivered exogenously to specific tissue or cellular targets as required. Alternatively, the nucleic acid molecules (*e.g.*, ribozymes and antisense) can be expressed from DNA and/or RNA vectors that are delivered to specific cells.

The enzymatic nucleic acid-based inhibitors of HBV expression are useful for the prevention of the diseases and conditions including HBV infection, hepatitis, cancer, cirrhosis, liver failure, and any other diseases or conditions that are related to the levels of HBV in a cell or tissue.

The nucleic acid-based inhibitors of the invention are added directly, or can be complexed with cationic lipids, packaged within liposomes, or otherwise delivered to target cells or tissues. The nucleic acid or nucleic acid complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo* through injection, infusion pump or stent, with or without their incorporation in biopolymers. In preferred embodiments, the enzymatic nucleic acid HBV inhibitors comprise sequences, which are complementary to the substrate sequences in. Examples of such enzymatic nucleic acid molecules also are shown in. Examples of such enzymatic nucleic acid molecules consist essentially of sequences defined in these tables.

In yet another embodiment, the invention features antisense nucleic acid molecules including sequences complementary to the HBV substrate sequences shown in. Such nucleic acid molecules can include sequences as shown for the binding arms of the enzymatic nucleic acid molecules in. Similarly, triplex molecules can be provided targeted to the corresponding DNA target regions, and regions containing the DNA equivalent of a target sequence or a sequence complementary to the specified target (substrate) sequence. Typically, antisense molecules are complementary to a target sequence along a single contiguous sequence of the antisense molecule. However, in certain embodiments, an antisense molecule can bind to substrate such that the substrate molecule forms a loop, and/or an antisense molecule can bind such that the antisense molecule forms a loop. Thus, the antisense molecule can be complementary to two (or even more) non-contiguous substrate sequences or two (or even more) non-contiguous sequence portions of an antisense molecule can be complementary to a target sequence or both.

By "consists essentially of" is meant that the active nucleic acid molecule of the invention, for example, an enzymatic nucleic acid molecule, contains an enzymatic center or core equivalent to those in the examples, and binding arms able to bind RNA such that cleavage at the target site occurs. Other sequences can be present which do not interfere with such cleavage. Thus, a core region can, for example, include one or more loops, stem-loop structure, or linker which does not prevent enzymatic activity. Thus, the underlined regions in the sequences in can be such a loop, stem-loop, nucleotide linker, and/or non-nucleotide linker and can be represented generally as sequence "X". For example, a core sequence for a hammerhead enzymatic nucleic acid can comprise a conserved sequence, such as 5'-CUGAUGAG-3' and 5'-CGAA-3' connected by "X", where X is 5'-<u>GCCGUUAGGC</u>-3' (SEQ ID NO. 16201), or any other Stem II region known in the art, or a nucleotide and/or non-nucleotide linker. Similarly, for other nucleic acid molecules of the instant invention, such as Inozyme, G-cleaver, amberzyme, zinzyme, DNAzyme, antisense, 2-5A antisense, triplex forming nucleic acid, and decoy nucleic acids, other sequences or non-nucleotide linkers can be present that do not interfere with the function of the nucleic acid molecule.

In another aspect of the invention, enzymatic nucleic acids or antisense molecules that interact with target RNA molecules and inhibit HBV (specifically HBV RNA) activity are expressed from transcription units inserted into DNA or RNA vectors. The recombinant vectors are preferably DNA plasmids or viral vectors. Enzymatic nucleic acid or antisense expressing viral vectors can be constructed based on, but not limited to, adeno-associated virus, retrovirus, adenovirus, or alphavirus. Preferably, the recombinant vectors capable of expressing the enzymatic nucleic acids or antisense are delivered as described above, and persist in target cells. Alternatively, viral vectors can be used that provide for transient expression of enzymatic nucleic acids or antisense. Such vectors can be repeatedly administered as necessary. Once expressed, the enzymatic nucleic acids or antisense bind to the target RNA and inhibit its function or expression. Delivery of enzymatic nucleic acids or antisense expressing vectors can be systemic, such as by intravenous or intramuscular administration, by administration to target cells ex-planted from the patient followed by reintroduction into the patient, or by any other means that allow for introduction into the desired target cell. Antisense DNA can be expressed via the use of a single stranded DNA intracellular expression vector.

In another embodiment, the invention features nucleic acid-based inhibitors (*e.g.*, enzymatic nucleic acid molecules (ribozymes), antisense nucleic acids, triplex DNA, decoys, aptamers, siRNA, antisense nucleic acids containing RNA cleaving chemical groups) and methods for their use to down regulate or inhibit the expression of RNA (*e.g.*, HBV) capable of progression and/or maintenance of liver disease and failure.

In another embodiment, the invention features nucleic acid-based techniques (*e.g.*, enzymatic nucleic acid molecules (ribozymes), antisense nucleic acids, triplex DNA, decoys, aptamers, siRNA, antisense nucleic acids containing RNA cleaving chemical groups) and methods for their use to down regulate or inhibit the expression of HBV RNA expression.

In other embodiments, the invention features a method for the analysis of HBV proteins. This method is useful in determining the efficacy of HBV inhibitors. Specifically, the instant invention features an assay for the analysis of HBsAg proteins and secreted alkaline phosphatase (SEAP) control proteins to determine the efficacy of agents used to modulate HBV expression.

The method consists of coating a micro-titer plate with an antibody such as anti-HBsAg Mab (for example, Biostride B88-95-31ad,ay) at 0.1 to 10 μg/ml in a buffer (for example, carbonate buffer, such as Na_2CO_3 15 mM, $NaHCO_3$ 35 mM, pH 9.5) at 4°C overnight. The microtiter wells are then washed with PBST or the equivalent thereof, (for example, PBS, 0.05% Tween 20) and blocked for 0.1-24 hr at 37° C with PBST, 1% BSA or the equivalent thereof. Following washing as above, the wells are dried (for example, at 37° C for 30 min).

Biotinylated goat anti-HBsAg or an equivalent antibody (for example, Accurate YVS1807) is diluted (for example at 1:1000) in PBST and incubated in the wells (for example, 1 hr. at 37° C). The wells are washed with PBST (for example, 4x). A conjugate, (for example, Streptavidin/Alkaline Phosphatase Conjugate, Pierce 21324) is diluted to 10-10,000 ng/ml in PBST, and incubated in the wells (for example, 1 hr. at 37° C). After washing as above, a substrate (for example, p-nitrophenyl phosphate substrate, Pierce 37620) is added to the wells, which are then incubated (for example, 1 hr. at 37° C). The optical density is then determined (for example, at 405 nm). SEAP levels are then assayed, for example, using the Great EscAPe® Detection Kit (Clontech K2041-1), as per the manufacturers instructions. In the above example, incubation times and reagent concentrations can be varied to achieve optimum results, a non-limiting example is described in Example 6.

Comparison of this HBsAg ELISA method to a commercially available assay from World Diagnostics, Inc. 15271 NW 60th Ave, #201, Miami Lakes, FL 33014 (305) 827-3304 (Cat. No. EL10018) demonstrates an increase in sensitivity (signal:noise) of 3-20 fold.

This invention also relates to nucleic acid molecules directed to disrupt the function of HBV reverse transcriptase. In addition, the invention relates to nucleic acid molecules directed to disrupt the function of the Enhancer I core region of the HBV genomic DNA. In particular, the present invention describes the selection and function of nucleic acid molecules, such as decoys and aptamers, capable of specifically binding to the HBV reverse transcriptase (pol) primer and modulating reverse transcription of the HBV pregenomic RNA. In another embodiment, the present invention relates to nucleic acid molecules, such as decoys, antisense and aptamers, capable of specifically binding to the HBV reverse transcriptase (pol) and modulating reverse transcription of the HBV pregenomic RNA. In yet another embodiment, the present invention relates to nucleic acid molecules capable of specifically binding to the HBV Enhancer I core region and modulating transcription of the HBV genomic DNA. The invention further relates to allosteric enzymatic nucleic acid molecules or "allozymes" that are used to modulate HBV gene expression. Such allozymes are active in the presence of HBV-derived nucleic acids, peptides, and/or proteins such as HBV reverse transcriptase and/or a HBV reverse transcriptase primer sequence, thereby allowing the allozyme to selectively cleave a sequence of HBV DNA or RNA. Allozymes of the invention are also designed to be active in the presence of HBV Enhancer I sequences and/or mutant HBV Enhancer I sequences, thereby allowing the allozyme to selectively cleave a sequence of HBV DNA or RNA. These nucleic acid molecules can be used to treat diseases and disorders associated with HBV infection.

In one embodiment, the invention features a nucleic acid decoy molecule that specifically binds the hepatitis B virus (HBV) reverse transcriptase primer sequence. In

another embodiment, the invention features a nucleic acid decoy molecule that specifically binds the hepatitis B virus (HBV) reverse transcriptase. In yet another embodiment, the invention features a nucleic acid decoy molecule that specifically binds to the HBV Enhancer I core sequence.

In one embodiment, the invention features a nucleic acid aptamer that specifically binds the hepatitis B virus (HBV) reverse transcriptase primer. In another embodiment, the invention features a nucleic acid aptamer that specifically binds the hepatitis B virus (HBV) reverse transcriptase. In yet another embodiment, the invention features a nucleic acid aptamer molecule that specifically binds to the HBV Enhancer I core sequence.

In one embodiment, the invention features an allozyme that specifically binds the hepatitis B virus (HBV) reverse transcriptase primer. In another embodiment, the invention features an allozyme that specifically binds the hepatitis B virus (HBV) reverse transcriptase. In yet another embodiment, the invention features an allozyme that specifically binds to the HBV Enhancer I core sequence.

In yet another embodiment, the invention features a nucleic acid molecule, for example a triplex forming nucleic acid molecule or antisense nucleic acid molecule, that binds the hepatitis B virus (HBV) reverse transcriptase primer. In another embodiment, the invention features a triplex forming nucleic acid molecule or antisense nucleic acid molecule that specifically binds the hepatitis B virus (HBV) reverse transcriptase. In yet another embodiment, the invention features a triplex forming nucleic acid molecule or antisense nucleic acid molecule that specifically binds to the HBV Enhancer I core sequence.

In another embodiment, a nucleic acid molecule of the invention binds to Hepatocyte Nuclear Factor 3 (HNF3) and/or Hepatocyte Nuclear Factor 4 (HNF4) binding sequence within the HBV Enhancer I region of HBV genomic DNA, for example the plus strand and/or minus strand DNA of the Enhancer I region, and blocks the binding of HNF3 and/or HNF4 to the Enhancer 1 region.

In another embodiment, the nucleic acid molecule of the invention comprises a sequence having $(UUCA)_n$ domain, where n is an integer from 1-10. In another embodiment, the nucleic acid molecules of the invention comprise the sequence of SEQ. ID NOs: 11216 - 11342.

In another embodiment, the invention features a composition comprising a nucleic acid molecule of the invention and a pharmaceutically acceptable carrier. In another embodiment, the invention features a mammalian cell, for example a human cell, including a nucleic acid molecule contemplated by the invention.

In one embodiment, the invention features a method for treatment of HBV infection, cirrhosis, liver failure, or hepatocellular carcinoma, comprising administering to a patient a nucleic acid molecule of the invention under conditions suitable for the treatment.

In another embodiment, the invention features a method for the treatment of a patient having a condition associated with HBV infection comprising contacting cells of said patient with a nucleic acid molecule of the invention under conditions suitable for such treatment. In another embodiment, the invention features a method for the treatment of a patient having a condition associated with HBV infection comprising contacting cells of said patient with a nucleic acid molecule of the invention, and further comprising the use of one or more drug therapies, for example type I interferon or 3TC® (lamivudine), under conditions suitable for said treatment. In another embodiment, the other therapy is administered simultaneously with or separately from the nucleic acid molecule.

In another embodiment, the invention features a method for modulating HBV replication in a mammalian cell comprising administering to the cell a nucleic acid molecule of the invention under conditions suitable for the modulation.

In yet another embodiment, the invention features a method of modulating HBV reverse transcriptase activity comprising contacting a nucleic acid molecule of the invention, for example a decoy or aptamer, with HBV reverse transcriptase under conditions suitable for the modulating of the HBV reverse transcriptase activity.

In another embodiment, the invention features a method of modulating HBV transcription comprising contacting a nucleic molecule of the invention with a HBV Enhancer I sequence under conditions suitable for the modulation of HBV transcription.

In one embodiment, a nucleic acid molecule of the invention, for example a decoy or aptamer, is chemically synthesized. In another embodiment, the nucleic acid molecule of the invention comprises at least one nucleic acid sugar modification. In yet another embodiment, the nucleic acid molecule of the invention comprises at least one nucleic acid base modification. In another embodiment, the nucleic acid molecule of the invention comprises at least one nucleic acid backbone modification.

In another embodiment, the nucleic acid molecule of the invention comprises at least one 2'-O-alkyl, 2'-alkyl, 2'-alkoxylalkyl, 2'-alkylthioalkyl, 2'-amino, 2'-O-amino, or 2'-halo modification and/or any combination thereof with or without 2'-deoxy and/or 2'-ribo nucleotides. In yet another embodiment, the nucleic acid molecule of the invention comprises all 2'-O-alkyl nucleotides, for example, all 2'-O-allyl nucleotides.

In one embodiment, the nucleic acid molecule of the invention comprises a 5'-cap, 3'-cap, or 5'-3' cap structure, for example an abasic or inverted abasic moiety.

In another embodiment, the nucleic acid molecule of the invention is a linear nucleic acid molecule. In another embodiment, the nucleic acid molecule of the invention is a linear nucleic acid molecule that can optionally form a hairpin, loop, stem-loop, or other secondary structure. In yet another embodiment, the nucleic acid molecule of the invention is a circular nucleic acid molecule.

In one embodiment, the nucleic acid molecule of the invention is a single stranded oligonucleotide. In another embodiment, the nucleic acid molecule of the invention is a double-stranded oligonucleotide.

In one embodiment, the nucleic acid molecule of the invention comprises an oligonucleotide having between about 3 and about 100 nucleotides. In another embodiment, the nucleic acid molecule of the invention comprises an oligonucleotide having between about 3 and about 24 nucleotides. In another embodiment, the nucleic acid molecule of the invention comprises an oligonucleotide having between about 4 and about 16 nucleotides.

The nucleic acid decoy molecules and/or aptamers that bind to a reverse transcriptase and/or reverse transcriptase primer and therefore inactivate the reverse transcriptase, represent a novel therapeutic approach to treat a variety of pathologic indications, including, viral infection such as HBV infection, hepatitis, hepatocellular carcinoma, tumorigenesis, cirrhosis, liver failure and others.

The nucleic acid molecules that bind to a HBV Enhancer I sequence and therefore inactivate HBV transcription, represent a novel therapeutic approach to treat a variety of pathologic indications, including viral infection such as HBV infection, hepatitis, hepatocellular carcinoma, tumorigenesis, cirrhosis, liver failure and others conditions associated with the level of HBV.

In one embodiment of the present invention, a decoy nucleic acid molecule of the invention is 4 to 50 nucleotides in length, in specific embodiments about 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 nucleotides in length. In another embodiment, a non-decoy nucleic acid molecule, *e.g.*, an antisense molecule, a triplex DNA, or a ribozyme, is 13 to 100 nucleotides in length, *e.g.*, in specific embodiments 35, 36, 37, or 38 nucleotides in length (*e.g.*, for particular ribozymes or antisense). In particular embodiments, the nucleic acid molecule is 15-100, 17-100, 20-100, 21-100, 23-100, 25-100, 27-100, 30-100, 32-100, 35-100, 40-100, 50-100, 60-100, 70-100, or 80-100 nucleotides in length. Instead of 100 nucleotides being the upper limit on the length ranges specified above, the upper limit of the

length range can be, for example, 30, 40, 50, 60, 70, or 80 nucleotides. Thus, for any of the length ranges, the length range for particular embodiments has lower limit as specified, with an upper limit as specified which is greater than the lower limit. For example, in a particular embodiment, the length range can be 35-50 nucleotides in length. All such ranges are expressly included. Also in particular embodiments, a nucleic acid molecule can have a length which is any of the lengths specified above, for example, 21 nucleotides in length.

Exemplary nucleic acid decoy molecules of the invention are shown in **Table XIV**. Exemplary synthetic nucleic acid molecules of the invention are shown in **Table XV**. For example, decoy molecules of the invention are between 4 and 40 nucleotides in length. Exemplary decoys of the invention are 4, 8, 12, or 16 nucleotides in length. In an additional example, enzymatic nucleic acid molecules of the invention are preferably between 15 and 50 nucleotides in length, more preferably between 25 and 40 nucleotides in length, *e.g.*, 34, 36, or 38 nucleotides in length (for example see Jarvis *et al.*, 1996, *J. Biol. Chem.*, 271, 29107-29112). Exemplary DNAzymes of the invention are preferably between 15 and 40 nucleotides in length, more preferably between 25 and 35 nucleotides in length, *e.g.*, 29, 30, 31, or 32 nucleotides in length (see for example Santoro *et al.*, 1998, *Biochemistry*, 37, 13330-13342; Chartrand *et al.*, 1995, Nucleic Acids Research, 23, 4092-4096). Exemplary antisense molecules of the invention are preferably between 15 and 75 nucleotides in length, more preferably between 20 and 35 nucleotides in length, *e.g.*, 25, 26, 27, or 28 nucleotides in length (see for example Woolf *et al.*, 1992, *PNAS.*, 89, 7305-7309; Milner *et al.*, 1997, *Nature Biotechnology*, 15, 537-541). Exemplary triplex forming oligonucleotide molecules of the invention are preferably between 10 and 40 nucleotides in length, more preferably between 12 and 25 nucleotides in length, *e.g.*, 18, 19, 20, or 21 nucleotides in length (see for example Maher *et al.*, 1990, *Biochemistry*, 29, 8820-8826; Strobel and Dervan, 1990, *Science*, 249, 73-75). Those skilled in the art will recognize that all that is required is that the nucleic acid molecule is of length and conformation sufficient and suitable for the nucleic acid molecule to catalyze a reaction contemplated herein. The length of the nucleic acid molecules of the instant invention are not limiting within the general limits stated.

In one embodiment, the invention provides a method for producing a class of nucleic acid–based gene modulating agents, which exhibit a high degree of specificity for a viral reverse transcriptase such as HBV reverse transcriptase or reverse transcriptase primer such as a HBV reverse transcriptase primer. For example, the nucleic acid molecule is preferably targeted to a highly conserved nucleic acid binding region of the viral reverse transcriptase such that specific treatment of a disease or condition can be provided with either one or several nucleic acid molecules of the invention. Such nucleic acid molecules can be delivered exogenously to specific tissue or cellular targets as required. Alternatively, the

nucleic acid molecules can be expressed from DNA and/or RNA vectors that are delivered to specific cells.

In another embodiment, the invention provides a method for producing a class of nucleic acid–based gene modulating agents which exhibit a high degree of specificity for a viral enhancer regions such as the HBV Enhancer I core sequence. For example, the nucleic acid molecule is preferably targeted to a highly conserved transcription factor-binding region of the viral Enhancer I sequence such that specific treatment of a disease or condition can be provided with either one or several nucleic acid molecules of the invention. Such nucleic acid molecules can be delivered exogenously to specific tissue or cellular targets as required. Alternatively, the nucleic acid molecules can be expressed from DNA and/or RNA vectors that are delivered to specific cells.

In a another embodiment the invention provides a method for producing a class of enzymatic cleaving agents which exhibit a high degree of specificity for the RNA of a desired target. The enzymatic nucleic acid molecule, nuclease activating compound or chimera is preferably targeted to a highly conserved sequence region of a target mRNAs encoding HCV or HBV proteins such that specific treatment of a disease or condition can be provided with either one or several enzymatic nucleic acids. Such nucleic acid molecules can be delivered exogenously to specific cells as required. Alternatively, the enzymatic nucleic acid molecules can be expressed from DNA/RNA vectors that are delivered to specific cells. DNAzymes can be synthesized chemically or expressed endogenously *in vivo*, by means of a single stranded DNA vector or equivalent thereof.

In another embodiment, the nucleic acid molecule of the invention binds irreversibly to the HBV reverse transcriptase target, for example by covalent attachment of the nucleic molecule to the reverse transcriptase primer sequence. The covalent attachment can be accomplished by introducing chemical modifications into the nucleic acid molecule's (for example, decoy or aptamer) sequence that are capable of forming covalent bonds to the reverse transcriptase primer sequence.

In another embodiment, the nucleic acid molecule of the invention binds irreversibly to the HBV Enhancer I sequence target, for example, by covalent attachment of the nucleic acid molecule to the HBV Enhancer I sequence. The covalent attachment can be accomplished by introducing chemical modifications into the nucleic acid molecule's sequence that are capable of forming covalent bonds to the reverse transcriptase primer sequence.

In another embodiment, the type I interferon contemplated by the invention is interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon,

polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, polyethylene glycol consensus interferon.

In one embodiment, the invention features a composition comprising type I interferon and a nucleic acid molecule of the inventionand a pharmaceutically acceptable carrier.

In another embodiment, the invention features a method of administering to a cell, for example a mammalian cell or human cell, a nucleic acid molecule of the invention independently or in conjunction with other therapeutic compounds, such as type I interferon or 3TC® (lamivudine), comprising contacting the cell with the nucleic acid molecule under conditions suitable for the administration.

In yet another embodiment, the invention features a method of administering to a cell, for example a mammalian cell or human cell, a nucleic acid molecule of the invention independently or in conjunction with other therapeutic compounds such as enzymatic nucleic acid molecules, antisense molecules, triplex forming oligonucleotides, 2,5-A chimeras, and/or RNAi, comprising contacting the cell with the nucleic acid molecule of the invention under conditions suitable for the administration.

In another embodiment, administration of a nucleic acid molecule of the invention is administered to a cell or patient in the presence of a delivery reagent, for example a lipid, cationic lipid, phospholipid, or liposome.

In one embodiment, the invention features novel nucleic acid-based techniques such as nucleic acid decoy molecules and/or aptamers, used alone or in combination with enzymatic nucleic acid molecules, antisense molecules, and/or RNAi, and methods for use to down regulate or modulate the expression of HBV RNA and/or replication of HBV.

In another embodiment, the invention features the use of one or more of the nucleic acid-based techniques to modulate the expression of the genes encoding HBV viral proteins. Specifically, the invention features the use of nucleic acid-based techniques to specifically modulate the expression of the HBV viral genome.

In another embodiment, the invention features the use of one or more of the nucleic acid-based techniques to modulate the activity, expression, or level of cellular proteins required for HBV replication. For example, the invention features the use of nucleic acid-based techniques to specifically modulate the activity of cellular proteins required for HBV replication.

In another embodiment, the invention features nucleic acid-based modulators(*e.g.*, nucleic acid decoy molecules, aptamers, enzymatic nucleic acid molecules (ribozymes),

antisense nucleic acids, triplex DNA, antisense nucleic acids containing RNA cleaving chemical groups) and methods for their use to down regulate or modulate reverse transcriptase activity and/or the expression of RNA (*e.g.,* HBV) capable of progression and/or maintenance of HBV infection, hepatocellular carcinoma, liver disease and failure.

In another embodiment, the invention features nucleic acid-based techniques (*e.g.,* nucleic acid decoy molecules, aptamers, enzymatic nuleic acid molecules (ribozymes), antisense nucleic acid molecules, triplex DNA, antisense nucleic acids containing RNA cleaving chemical groups) and methods for their use to down regulate or modulate reverse transcriptase activity and/or the expression of HBV RNA.

In another embodiment, the invention features nucleic acid-based modulators (*e.g.,* nucleic acid decoy molecules, aptamers, enzymatic nucleic acid molecules (ribozymes), antisense nucleic acids, triplex DNA, siRNA, dsRNA, antisense nucleic acids containing RNA cleaving chemical groups) and methods for their use to down regulate or modulate Enhancer I mediated transcription activity and/or the expression of DNA (*e.g.,* HBV) capable of progression and/or maintenance of HBV infection, hepatocellular carcinoma, liver disease and failure.

In another embodiment, the invention features nucleic acid-based techniques (*e.g.,* nucleic acid decoy molecules, aptamers, enzymatic nucleic acid molecules, antisense nucleic acid molecules, triplex DNA, siRNA, antisense nucleic acids containing DNA cleaving chemical groups) and methods for their use to down regulate or modulate Enhancer I mediated transcription activity and/or the expression of HBV DNA.

In another embodiment, the invention features a nucleic acid sensor molecule having an enzymatic nucleic acid domain and a sensor domain that interacts with an HBV peptide, protein, or polynucleotide sequence, for example, HBV reverse transcriptase, HBV reverse transcriptase primer, or the Enhancer I element of the HBV pregenomic RNA, wherein such interaction results in modulation of the activity of the enzymatic nucleic acid domain of the nucleic acid sensor molecule. In another embodiment, the invention features HBV-specific nucleic acid sensor molecules or allozymes, and methods for their use to down regulate or modulate the expression of HBV RNA capable of progression and/or maintenance of hepatitis, hepatocellular carcinoma, cirrhosis, and/or liver failure. In yet another embodiment, the enzymatic nucleic acid domain of a nucleic acid sensor molecule of the invention is a Hammerhead, Inozyme, G-cleaver, DNAzyme, Zinzyme, Amberzyme, or Hairpin enzymatic nucleic acid molecule.

In one embodiment, nucleic acid molecules of the invention are used to treat HBV-infected cells or a HBV-infected patient wherein the HBV is resistant or the patient does not

respond to treatment with 3TC® (Lamivudine), either alone or in combination with other therapies under conditions suitable for the treatment.

In another embodiment, nucleic acid molecules of the invention are used to treat HBV-infected cells or a HBV-infected patient, wherein the HBV is resistant or the patient does not respond to treatment with Interferon, for example Infergen®, either alone or in combination with other therapies under conditions suitable for the treatment.

The invention also relates to *in vitro* and *in vivo* systems, including, e.g., mammalian systems for screening inhibitors of HBV. In one embodiment, the invention features a mouse, for example a male or female mouse, implanted with HepG2.2.15 cells, wherein the mouse is susceptible to HBV infection and capable of sustaining HBV DNA expression. One embodiment of the invention provides a mouse implanted with HepG2.2.15 cells, wherein said mouse sustains the propagation of HEPG2.2.15 cells and HBV production.

In another embodiment, a mouse of the invention has been infected with HBV for at least one week to at least eight weeks, including, for example at least 4 weeks.

In yet another embodiment, a mouse of the invention, for example a male or female mouse, is an immunocompromised mouse, for example a nu/nu mouse or a scid/scid mouse.

In one embodiment, the invention features a method of producing a mouse of the invention, comprising injecting, for example by subcutaneous injection, HepG2.2.15 (Sells, *et al*,. 1987, *Proc Natl Acad Sci U S A.*, 84, 1005-1009) cells into the mouse under conditions suitable for the propagation of HepG2.2.15 cells in said mouse. HepG2.2.15 cells can be suspended in, for example, Delbecco's PBS solution including calcium and magnesium. In another embodiment, HepG2.2.15 cells are selected for antibiotic resistance and are then introduced into the mouse under conditions suitable for the propagation of HepG2.2.15 cells in said mouse. A non-limiting example of antibiotic resistant HepG2.2.15 cells include G418 antibiotic resistant HepG2.2.15 cells.

In another embodiment, the invention features a method of screening a compound for therapeutic activity against HBV, comprising administering the compound to a mouse of the invention and monitoring the the levels of HBV produced (e.g. by assaying for HBV DNA levels) in the mouse.

In one embodiment, a therapeutic compound or therapy contemplated by the invention is a lipid, steroid, peptide, protein, antibody, monoclonal antibody, humanized monoclonal antibody, small molecule, and/or isomers and analogs thereof, and/or a cell.

In one embodiment, a therapeutic compound or therapy contemplated by the invention is a nucleic acid molecule, for example a nucleic acid molecule, such as an enzymatic nucleic acid molecule, antisense nucleic acid molecule, allozyme, peptide nucleic acid, decoy, triplex oligonucleotide, dsRNA, ssRNA, RNAi, siRNA, aptamer, or 2,5-A chimera used alone or in combination with another therapy, for example antiviral therapy. Antiviral therapy can be, for example, treatment with 3TC® (Lamivudine) or interferon. Interferon can include, for example, consensus interferon or type I interferon. Type I interferon can include interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, or polyethylene glycol consensus interferon.

In one embodiment, the invention features a non-human mammal implanted with HepG2.2.15 cells, wherein the non-human mammal is susceptible to HBV infection and capable of sustaining HBV DNA expression in the implanted HepG2.2.15 cells.

In another embodiment, a non-human mammal of the invention, for example a male or female non-human mammal, has been infected with HBV for at least one week to at least eight weeks, including for example at least four weeks.

In yet another embodiment, a non-human mammal of the invention is an immunocompromised mammal, for example a nu/nu mammal or a scid/scid mammal.

In one embodiment, the invention features a method of producing a non-human mammal comprising HepG2.2.15 cells comprising injecting, for example by subcutaneous injection, HepG2.2.15 cells into the non-human mammal under conditions suitable for the propagation of HepG2.2.15 cells in said non-human mammal.

In another embodiment, the invention features a method of screening a compound for therapeutic activity against HBV comprising administering the compound to a non-human mammal of the invention and monitoring the levels of HBV produced (e.g. by assaying for HBV DNA levels) in the non-human mammals.

In one embodiment, a therapeutic compound or therapy contemplated by the invention is a nucleic acid molecule, for example an enzymatic nucleic acid molecule, allozyme, antisense nucleic acid molecule, decoy, triplex oligonucleotide, dsRNA, ssRNA, RNAi, siRNA, or 2,5-A chimera used alone or in combination with another therapy, for example antiviral therapy.

Methods and chimeric immunocompromised heterologous non-human mammalian hosts, particularly mouse hosts, are provided for the expression of hepatitis B virus ("HBV").

In one embodiment, the chimeric hosts have transplanted viable, HepG2.2.15 cells in an immunocompromised host.

The non-human mammals contemplated by the invention are immunocompromised in normally inheriting the desired immune incapacity, or the desired immune incapacity can be created. For example, hosts with severe combined immunodeficiency, known as scid/scid hosts, are available. Rodentia, particularly mice, and equine, particularly horses, are presently available as scid/scid hosts, for example scid/scid mice and scid/scid rats. The scid/scid hosts lack functioning lymphocyte types, particularly B-cells and some T-cell types. In the scid/scid mouse hosts, the genetic defect appears to be a non-functioning recombinase, as the germline DNA is not rearranged to produce functioning surface immunoglobulin and T-cell receptors.

Any immunodeficient non-human mammals, e.g. mouse, can be used to generate the animal models described herein. The term "immunodeficient," as used herein, refers to a genetic alteration that impairs the animal's ability to mount an effective immune response. In this regard, an "effective immune response" is one which is capable of destroying invading pathogens such as (but not limited to) viruses, bacteria, parasites, malignant cells, and/or a xenogeneic or allogeneic transplant. In one embodiment, the immunodeficient mouse is a severe immunodeficient (SCID) mouse, which lacks recombinase activity that is necessary for the generation of immunoglobulin and functional T cell antigen receptors, and thus does not produce functional B and T lymphocytes. In another embodiment, the immunodeficient mouse is a nude mouse, which contains a genetic defect that results in the absence of a functional thymus, leading to T-cell and B-cell deficiencies. However, mice containing other immunodeficiencies (such as rag-1 or rag-2 knockouts, as described in Chen *et al.,* 1994, *Curr. Opin. Immunol.,* 6, 313-319 and Guidas *et al.,* 1995, *J. Exp. Med.,* 181, 1187-1195, or beige-nude mice, which also lack natural killer cells, as described in Kollmann *et al.,* 1993, *J. Exp. Med.*, 177, 821-832) can also be employed.

The introduction of HepG2.2.15 cells occurs with a host at an age less than about 25% of its normal lifespan, usually to 20% of the normal lifespan with mice, and the age will generally be of an age of about 3 to 10 weeks, more usually from about 4 to 8 weeks. The mice can be of either sex, can be neutered, and can be otherwise normal, except for the immunocompromised state, or they can have one or more mutations, which can be naturally occurring or as a result of mutagenesis.

In another embodiment, the mouse model described herein is used to evaluate the effectiveness of thetherapeutic compounds and methods. The terms "therapeutic compounds", "therapeutic methods" and "therapy" as used herein, encompass exogenous factors, such as dietary or environmental conditions, as well as pharmaceutical compositions

"drugs" and vaccines. In one embodiment, the therapeutic method is an immunotherapy, which can include the treatment of the HBV bearing animal with populations of HBV-reactive immune cells. The therapeutic method can also, or alternatively, be a gene therapy (i.e., a therapy that involves treatment of the HBV-bearing mouse with a cell population that has been manipulated to express one or more genes, the products of which can possess anti-viral activity), see for example The Development of Human Gene Therapy, Theodore Friedmann, Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1999. Therapeutic compounds of the invention can comprise a drug or composition with pharmaceutical activity that can be used to treat illness or disease. A therapeutic method can comprise the use of a plurality of compounds in a mixture or a distinct entity. Examples of such compounds include nucleosides, nucleic acids, nucleic acid chimeras, RNA and DNA oligonucleotides, peptide nucleic acids, enzymatic nucleic acid molecules, antisense nucleic acid molecules, decoys, triplex oligonucleotides, ssDNA, dsRNA, ssRNA, siRNA, 2,5-A chimeras, lipids, steroids, peptides, proteins, antibodies, monoclonal antibodies (see for example Hall, 1995, Science, 270, 915-916), small molecules, and/or isomers and analogs thereof.

The methods of this invention can be used to treat human hepatitis B virus infections, which include productive virus infection, latent or persistent virus infection, and HBV-induced hepatocyte transformation. The utility can be extended to other species of HBV that infect non-human animals where such infections are of veterinary importance.

Preferred binding sites of the nucleic acid molecules of the invention include, but are not limited, to the primer binding site on HBV reverse transcriptase, the primer binding sequences of the HBV RNA, and/or the HBV Enhancer I region of HBV DNA.

This invention further relates to nucleic acid molecules that target RNA species of hepatitis C virus (HCV) and/or encoded by the HCV. In one embodiment, applicant describes enzymatic nucleic acid molecules that specifically cleave HCV RNA and the selection and function thereof. The invention further relates to compounds and chimeric molecules comprising nuclease activating activity. The invention also relates to compositions and methods for the cleavage of RNA using these nuclease activating compounds and chimeras. Nucleic acid molecules, nuclease activating compounds and chimeras, and compostions and methods of the invention can be used to treat diseases associated with HCV infection.

Due to the high sequence variability of the HCV genome, selection of nucleic acid molecules and nuclease activating compounds and chimeras for broad therapeutic applications preferably involve the conserved regions of the HCV genome. Thus, in one embodiment the present invention describes nucleic acid molecules that cleave the conserved

regions of the HCV genome. The invention further describes compounds and chimeric molecules that activate cellular nucleases that cleave HCV RNA, including concerved regions of the HCV genome. Examples of conserved regions of the HCV genome include but are not limited to the 5'-Non Coding Region (NCR), the 5'-end of the core protein coding region, and the 3'- NCR. HCV genomic RNA contains an internal ribosome entry site (IRES) in the 5'-NCR which mediates translation independently of a 5'-cap structure (Wang *et al.*, 1993, *J. Virol.*, 67, 3338-44). The full-length sequence of the HCV RNA genome is heterologous among clinically isolated subtypes, of which there are at least 15 (Simmonds, 1995, *Hepatology,* 21, 570-583), however, the 5'-NCR sequence of HCV is highly conserved across all known subtypes, most likely to preserve the shared IRES mechanism (Okamoto *et al.*, 1991, *J. General Virol.*, 72, 2697-2704). In general, enzymatic nucleic acid molecules and nuclease activating compounds, and chimeras that cleave sites located in the 5' end of the HCV genome are expected to block translation while nucleic acid molecules and nuclease activating compounds, and chimeras that cleave sites located in the 3' end of the genome are expected to block RNA replication. Therefore, one nucleic acid molecule, compound, or chimera can be designed to cleave all the different isolates of HCV. Enzymatic nucleic acid molecules and nuclease activating compounds, and chimeras designed against conserved regions of various HCV isolates enable efficient inhibition of HCV replication in diverse patient populations and ensure the effectiveness of the nucleic acid molecules and nuclease activating compounds, and chimeras against HCV quasi species which evolve due to mutations in the non-conserved regions of the HCV genome.

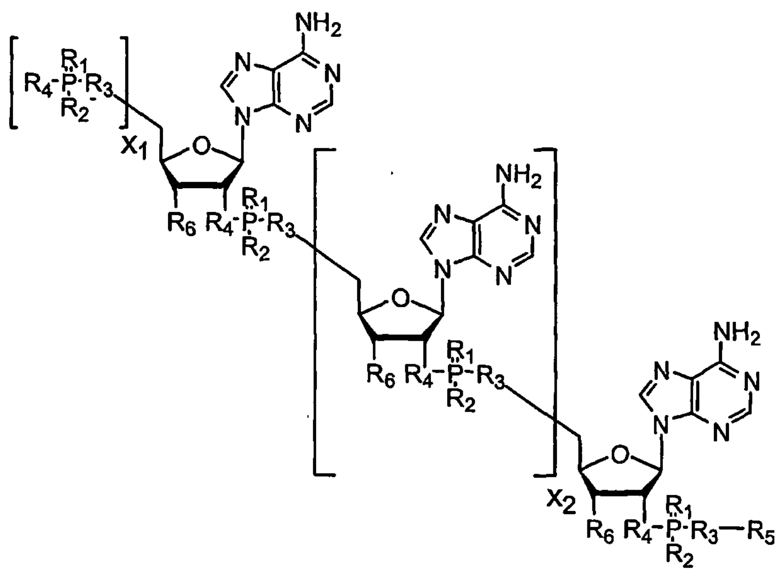
In one embodiment, the invention features an enzymatic nucleic acid molecule, preferably in the hammerhead, NCH (Inozyme), G-cleaver, amberzyme, zinzyme and/or DNAzyme motif, and the use thereof to down-regulate or inhibit the expression of HCV RNA.

In another embodiment, the invention features an enzymatic nucleic acid molecule, preferably in the hammerhead, Inozyme, G-cleaver, amberzyme, zinzyme and/or DNAzyme motif, and the use thereof to down-regulate or inhibit the expression of HCV minus strand RNA.

In yet another embodiment, the invention featues a nuclease activating compound and/or a chimera and the use thereof to down-regulate or inhibit the expression of HCV RNA.

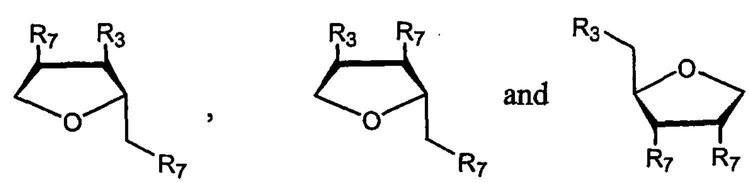
In another embodiment, the invention featues the use of a nuclease activating compound and/or a chimera to inhibit the expression of HCVminus strand RNA.

In one embodiment, the invention features a compound having formula I:



wherein X_1 is an integer selected from the group consisting of 1, 2, and 3; X_2 is an integer greater than or equal to 1; R_6 is independantly selected from the group including H, OH, NH_2, O NH_2, alkyl, S-alkyl, O-alkyl, O-alkyl-S-alkyl, O-alkoxyalkyl, allyl, O-allyl, and fluoro; each R_1 and R_2 are independantly selected from the group consisting of O and S; each R_3 and R_4 are independantly selected from the group consisting of O, N, and S; and R_5 is selected from the group consisting of alkyl, alkylamine, an oligonucleotide having any of SEQ ID NOS. 11343-16182, an oligonucleotide having a sequence complementary to a sequence selected from the group including SEQ ID NOS. 2594-7433, and abasic moiety.

In another embodiment, the abasic moiety of the instant invention is selected from the group consisting of:



wherein R_3 is selected from the group consisting of O, N, and S, and R_7 is independently selected from the group consisting of H, OH, NH2, O-NH2, alkyl, S-alkyl, O-alkyl, O-alkyl-S-alkyl, O-alkoxyalkyl, allyl, O-allyl, fluoro, oligonucleotide, alkyl, alkylamine and abasic moiety.

In another embodiment, the oligonucleotide R_5 of Formula I having a sequence complementary to a sequence selected from the group consisting of SEQ ID NOS. 2594-7433 is an enzymatic nucleic acid molecule.

In yet another embodiment, the oligonucleotide R_5 of Formula I having a sequence complementary to a sequence selected from the group consisting of SEQ ID NOS. 2594-7433 is an antisense nucleic acid molecule.

In another embodiment, the oligonucleotide R_5 of Formula I having a sequence complementary to a sequence selected from the group consisting of SEQ ID NOS. 2594-7433 is an enzymatic nucleic acid molecule selected from the group consisting of Hammerhead, Inozyme, G-cleaver, DNAzyme, Amberzyme, and Zinzyme motifs.

In another embodiment, the Inozyme enzymatic nucleic acid molecule of the instant invention comprises a stem II region of length greater than or equal to 2 base pairs.

In one embodiment, the oligonucleotide R_5 of Formula I having a sequence complementary to a sequence selected from the group consisting of SEQ ID NOS. 2594-7433 is an enzymatic nucleic acid comprising between 12 and 100 bases complementary to an RNA derived from HCV.

In another embodiment, the oligonucleotide R_5 of Formula I having a sequence complementary to a sequence selected from the group consisting of SEQ ID NOS. 2594-7433 is an enzymatic nucleic acid comprising between 14 and 24 bases complementary to said RNA derived from HCV.

In one embodiment, the oligonucleotide R_5 of Formula I having a sequence complementary to a sequence selected from the group consisting of SEQ ID NOS. 2594-7433 is an antisense nucleic acid comprising between 12 and 100 bases complementary to an RNA derived from HCV.

In another embodiment, the oligonucleotide R_5 of Formula I having a sequence complementary to a sequence selected from the group consisting of SEQ ID NOS. 2594-7433 is an antisense nucleic acid comprising between 14 and 24 bases complementary to said RNA derived from HCV.

In another embodiment, the invention features a composition comprising a compound of Formula I, in a pharmaceutically acceptable carrier.

In yet another embodiment, the invention features a mammalian cell comprising a compound of Formula I. For example, the mammalian cell comprising a compound of Formula I can be a human cell.

In one embodiment, the invention features a method for the treatment of cirrhosis, liver failure, hepatocellular carcinoma, or a condition associated with HCV infection comprising

the step of administering to a patient a compound of Formula I under conditions suitable for said treatment.

In another embodiment, the invention features a method of treatment of a patient having a condition associated with HCV infection comprising contacting cells of said patient with a compound having Formula I, and further comprising the use of one or more drug therapies under conditions suitable for said treatment. For example, the other therapies of the instant invention can be selected from the group consisting of type I interferon, interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, polyethylene glycol consensus interferon, treatment with an enzymatic nucleic acid molecule, and treatment with an antisense molecule.

In another embodiment, the other therapies of the instant invention, for example type I interferon, interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, polyethylene glycol consensus interferon, treatment with an enzymatic nucleic acid molecule, and treatment with an antisense nucleic acid molecule, and the compound having Formula I are administered separately in separate pharmaceutically acceptable carriers.

In yet another embodiment, the other therapies of the instant invention, for example type I interferon, interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, polyethylene glycol consensus interferon, treatment with an enzymatic nucleic acid molecule, and treatment with an antisense nucleic acid molecule, and the compound having Formula I are administered simultaneously in a pharmaceutically acceptable carrier. The invention features a composition comprising a compound of Formula I and one or more of the above-listed compounds in a pharmaceutically acceptable carrier.

In yet another embodiment, the invention features a method for inhibiting HCV replication in a mammalian cell comprising the step of administering to said cell a compound having Formula I under conditions suitable for said inhibition.

In another embodiment, the invention features a method of cleaving a separate RNA molecule (i.e., HCV RNA or RNA necessary for HCV replication) comprising contacting a compound having Formula I with the separate RNA molecule under conditions suitable for the cleavage of the separate RNA molecule. In one example, the method of cleaving a separate RNA molecule is carried out in the presence of a divalent cation, for example Mg2+.

In yet another embodiment, the method of cleaving a separate RNA molecule of the invention is carried out in the presence of a protein nuclease, for example RNAse L.

In one embodiment, a compound having Formula I is chemically synthesized. In one embodiment, a compound having Formula I comprises at least one 2'-sugar modification, at least one nucleic acid base modification, and/or at least one phosphate modification.

The nucleic acid-based modulators of the invention are added directly, or can be complexed with cationic lipids, packaged within liposomes, or otherwise delivered to target cells or tissues. The nucleic acid or nucleic acid complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo* through injection, infusion pump or stent, with or without their incorporation in biopolymers. In particular embodiments, the nucleic acid molecules of the invention comprise sequences shown in **Tables IV-XI, XIV-XV and XVIII-XXIII.** Examples of such nucleic acid molecules consist essentially of sequences defined in the tables.

The nucleic acid-based inhibitors, nuclease activating compounds and chimeras of the invention are added directly, or can be complexed with cationic lipids, packaged within liposomes, or otherwise delivered to target cells or tissues. The nucleic acid or nucleic acid complexes, and nuclease activating compounds or chimeras can be locally administered to relevant tissues ex vivo, or in vivo through injection or infusion pump, with or without their incorporation in biopolymers. In preferred embodiments, the enzymatic nucleic acid inhibitors, and nuclease activating compounds or chimeras comprise sequences, which are complementary to the substrate sequences in **Tables XVIII, XIX, XX and XXIII**. Examples of such enzymatic nucleic acid molecules also are shown in **Tables XVIII, XIX, XX, XXI and XXIII.** Examples of such enzymatic nucleic acid molecules consist essentially of sequences defined in these tables. In additional embodiments, the enzymatic nucleic acid inhibitors of the invention that comprise sequences which are complementary to the substrate sequences in **Tables XVIII, XIX, XX and XXIII** are covalently attached to nuclease activating compound or chimeras of the invention, for example a compound having Formula I.

In yet another embodiment, the invention features antisense nucleic acid molecules and 2-5A chimera including sequences complementary to the substrate sequences shown in **Tables XVIII, XIX, XX and XXIII.** Such nucleic acid molecules can include sequences as shown for the binding arms of the enzymatic nucleic acid molecules in **Tables XVIII, XIX, XX, XXI and XXIII.** Similarly, triplex molecules can be provided targeted to the corresponding DNA target regions, and containing the DNA equivalent of a target sequence or a sequence complementary to the specified target (substrate) sequence. Typically, antisense molecules are complementary to a target sequence along a single contiguous

sequence of the antisense molecule. However, in certain embodiments, an antisense molecule can bind to substrate such that the substrate molecule forms a loop, and/or an antisense molecule can bind such that the antisense molecule forms a loop. Thus, the antisense molecule can be complementary to two (or even more) non-contiguous substrate sequences or two (or even more) non-contiguous sequence portions of an antisense molecule can be complementary to a target sequence or both.

In one embodiment, the invention features nucleic acid molecules and nuclease activating compounds or chimeras that inhibit gene expression and/or viral replication. These chemically or enzymatically synthesized nucleic acid molecules can contain substrate binding domains that bind to accessible regions of their target mRNAs. The nucleic acid molecules also contain domains that catalyze the cleavage of RNA. The enzymatic nucleic acid molecules are preferably molecules of the hammerhead, Inozyme, DNAzyme, Zinzyme, Amberzyme, and/or G-cleaver motifs. Upon binding, the enzymatic nucleic acid molecules cleave the target mRNAs, preventing translation and protein accumulation. In the absence of the expression of the target gene, HCV gene expression and/or replication is inhibited.

In another aspect, the invention provides mammalian cells containing one or more nucleic acid molecules and/or expression vectors of this invention. The one or more nucleic acid molecules can independently be targeted to the same or different sites.

In one embodiment, nucleic acid decoys, aptamers, siRNA, enzymatic nucleic acids or antisense molecules that interact with target protein and/or RNA molecules and modulate HBV (specifically HBV reverse transcriptase, or transcription of HBV genomic DNA) activity are expressed from transcription units inserted into DNA or RNA vectors. The recombinant vectors are preferably DNA plasmids or viral vectors. Decoys, aptamers, enzymatic nucleic acid or antisense expressing viral vectors can be constructed based on, but not limited to, adeno-associated virus, retrovirus, adenovirus, or alphavirus. Preferably, the recombinant vectors capable of expressing the decoys, aptamers, enzymatic nucleic acids or antisense are delivered as described above, and persist in target cells. Alternatively, viral vectors can be used that provide for transient expression of decoys, aptamers, siRNA, enzymatic nucleic acids or antisense. Such vectors can be repeatedly administered as necessary. Once expressed, the decoys, aptamers, enzymatic nucleic acids or antisense bind to the target protein and/or RNA and modulate its function or expression. Delivery of decoy, aptamer, siRNA, enzymatic nucleic acid or antisense expressing vectors can be systemic, such as by intravenous or intramuscular administration, by administration to target cells ex-planted from the patient followed by reintroduction into the patient, or by any other means that would allow for introduction into the desired target cell. DNA based nucleic acid

molecules of the invention can be expressed via the use of a single stranded DNA intracellular expression vector.

In one embodiment, nucleic acid molecules and nuclease activating compounds or chimeras are added directly, or can be complexed with cationic lipids, packaged within liposomes, or otherwise delivered to target cells. The nucleic acid or nucleic acid complexes can be locally administered to relevant tissues ex vivo, or in vivo through injection, infusion pump or stent, with or without their incorporation in biopolymers. In another preferred embodiment, the nucleic acid molecule, nuclease activating compound or chimera is administered to the site of HBV or HCV activity (e.g., hepatocytes) in an appropriate liposomal vehicle.

In another embodiment, nucleic acid molecules that cleave target molecules and inhibit HCV activity are expressed from transcription units inserted into DNA or RNA vectors. The recombinant vectors are preferably DNA plasmids or viral vectors. Nucleic acid molecule expressing viral vectors can be constructed based on, but not limited to, adeno-associated virus, retrovirus, adenovirus, or alphavirus. Preferably, the recombinant vectors capable of expressing the nucleic acid molecules are delivered as described above, and persist in target cells. Alternatively, viral vectors can be used that provide for transient expression of nucleic acid molecules. Such vectors can be repeatedly administered as necessary. Once expressed, the nucleic acid molecules cleave the target mRNA. Delivery of enzymatic nucleic acid molecule expressing vectors can be systemic, such as by intravenous or intramuscular administration, by administration to target cells ex-planted from the patient followed by reintroduction into the patient, or by any other means that would allow for introduction into the desired target cell (for a review see Couture and Stinchcomb, 1996, *TIG.*, 12, 510). In another aspect of the invention, nucleic acid molecules that cleave target molecules and inhibit viral replication are expressed from transcription units inserted into DNA, RNA, or viral vectors. Preferably, the recombinant vectors capable of expressing the nucleic acid molecules are locally delivered as described above, and transiently persist in smooth muscle cells. However, other mammalian cell vectors that direct the expression of RNA can be used for this purpose.

The nucleic acid molecules of the instant invention, individually, or in combination or in conjunction with other drugs, and/or therapies can be used to treat diseases or conditions discussed herein. For example, to treat a disease or condition associated with the levels of HBV or HCV, the nucleic acid molecules can be administered to a patient or can be administered to other appropriate cells evident to those skilled in the art, individually or in combination with one or more drugs under conditions suitable for the treatment.

In a further embodiment, the described molecules, such as decoys, aptamers, antisense, enzymatic nucleic acids, or nuclease activating compounds and chimeras can be used in combination with other known treatments to treat conditions or diseases discussed above. For example, the described molecules could be used in combination with one or more known therapeutic agents to treat HBV infection, HCV infection, hepatitis, hepatocellular carcinoma, cancer, cirrhosis, and liver failure. Such therapeutic agents can include, but are not limited to, nucleoside analogs selected from the group comprising Lamivudine (3TC®), L-FMAU, and/or adefovir dipivoxil (for a review of applicable nucleoside analogs, see Colacino and Staschke, 1998, *Progress in Drug Research*, 50, 259-322). Immunomodulators selected from the group comprising Type 1 Interferon, therapeutic vaccines, steriods, and 2'-5' oligoadenylates (for a review of 2'-5' Oligoadenylates, see Charubala and Pfleiderer, 1994, *Progress in Molecular and Subcellular Biology*, 14, 113-138).

Nucleic acid molecules, nuclease activating compounds and chimeras of the invention, individually, or in combination or in conjunction with other drugs, can be used to treat diseases or conditions discussed above. For example, to treat a disease or condition associated with HBV or HCV levels, the patient can be treated, or other appropriate cells can be treated, as is evident to those skilled in the art.

In a further embodiment, the described molecules can be used in combination with other known treatments to treat conditions or diseases discussed above. For example, the described molecules can be used in combination with one or more known therapeutic agents to treat liver failure, hepatocellular carcinoma, cirrhosis, and/or other disease states associated with HBV or HCV infection. Additional known therapeutic agents are those comprising antivirals, interferons, and/or antisense compounds.

The term "inhibit" or "down-regulate" as used herein refers to the expression of the gene, or level of RNAs or equivalent RNAs encoding one or more protein subunits or components, or activity of one or more protein subunits or components, such as HBV protein or proteins, is reduced below that observed in the absence of the therapies of the invention. In one embodiment, inhibition or down-regulation with enzymatic nucleic acid molecule preferably is below that level observed in the presence of an enzymatically inactive or attenuated molecule that is able to bind to the same site on the target RNA, but is unable to cleave that RNA. In another embodiment, inhibition or down-regulation with antisense oligonucleotides is preferably below that level observed in the presence of, for example, an oligonucleotide with scrambled sequence or with mismatches. In another embodiment, inhibition or down-regulation of HBV with the nucleic acid molecule of the instant invention is greater in the presence of the nucleic acid molecule than in its absence.

The term "up-regulate" as used herein refers to the expression of the gene, or level of RNAs or equivalent RNAs encoding one or more protein subunits or components, or activity of one or more protein subunits or components, such as HBV or HCV protein or proteins, is greater than that observed in the absence of the therapies of the invention. For example, the expression of a gene, such as HBV or HCV genes, can be increased in order to treat, prevent, ameliorate, or modulate a pathological condition caused or exacerbated by an absence or low level of gene expression.

The term "modulate" as used herein refers to the expression of the gene, or level of RNAs or equivalent RNAs encoding one or more protein subunits or components, or activity of one or more proteins is up-regulated or down-regulated, such that the expression, level, or activity is greater than or less than that observed in the absence of the therapies of the invention.

The term "decoy " as used herein refers to a nucleic acid molecule, for example RNA or DNA, or aptamer that is designed to preferentially bind to a predetermined ligand. Such binding can result in the inhibition or activation of a target molecule. A decoy or aptamer can compete with a naturally occurring binding target for the binding of a specific ligand. For example, it has been shown that over-expression of HIV trans-activation response (TAR) RNA can act as a "decoy" and efficiently binds HIV tat protein, thereby preventing it from binding to TAR sequences encoded in the HIV RNA (Sullenger *et al.*, 1990, Cell, 63, 601-608). This is but a specific example and those in the art will recognize that other embodiments can be readily generated using techniques generally known in the art, see for example Gold *et al.*, 1995, *Annu. Rev. Biochem.*, 64, 763; Brody and Gold, 2000, *J. Biotechnol.*, 74, 5; Sun, 2000, *Curr. Opin. Mol. Ther.*, 2, 100; Kusser, 2000, *J. Biotechnol.*, 74, 27; Hermann and Patel, 2000, *Science*, 287, 820; and Jayasena, 1999, *Clinical Chemistry*, 45, 1628. Similarly, a decoy can be designed to bind to HBV or HCV proteins and block the binding of HBV or HCV DNA or RNA or a decoy can be designed to bind to HBV or HCV proteins and prevent molecular interaction with the HBV or HCV proteins.

By "aptamer" or "nucleic acid aptamer" as used herein is meant a nucleic acid molecule that binds specifically to a target molecule wherein the nucleic acid molecule has sequence that is distinct from sequence recognized by the target molecule in its natural setting. Alternately, an aptamer can be a nucleic acid molecule that binds to a target molecule where the target molecule does not naturally bind to a nucleic acid. The target molecule can be any molecule of interest. For example, the aptamer can be used to bind to a ligand-binding domain of a protein, thereby preventing interaction of the naturally occurring ligand with the protein. This is a non-limiting example and those in the art will recognize that other embodiments can be readily generated using techniques generally known in the art, see for

example Gold *et al.*, 1995, *Annu. Rev. Biochem.*, 64, 763; Brody and Gold, 2000, *J. Biotechnol.*, 74, 5; Sun, 2000, *Curr. Opin. Mol. Ther.*, 2, 100; Kusser, 2000, *J. Biotechnol.*, 74, 27; Hermann and Patel, 2000, *Science*, 287, 820; and Jayasena, 1999, *Clinical Chemistry*, 45, 1628.

By "enzymatic nucleic acid molecule" is meant a nucleic acid molecule that has complementarity in a substrate binding region to a specified gene target, and also has an enzymatic activity which is active to specifically cleave a target RNA molecule. That is, the enzymatic nucleic acid molecule is able to intermolecularly cleave a RNA molecule and thereby inactivate a target RNA molecule. These complementary regions allow sufficient hybridization of the enzymatic nucleic acid molecule to a target RNA molecule and thus permit cleavage. One hundred percent complementarity is preferred, but complementarity as low as 50-75% may also be useful in this invention (see for example Werner and Uhlenbeck, 1995, *Nucleic Acids Research*, 23, 2092-2096; Hammann *et al.*, 1999, *Antisense and Nucleic Acid Drug Dev.*, 9, 25-31). The nucleic acids can be modified at the base, sugar, and/or phosphate groups. The term enzymatic nucleic acid is used interchangeably with phrases such as ribozymes, catalytic RNA, enzymatic RNA, catalytic DNA, aptazyme or aptamer-binding ribozyme, regulatable ribozyme, catalytic oligonucleotides, nucleozyme, DNAzyme, RNA enzyme, endoribonuclease, endonuclease, minizyme, leadzyme, oligozyme or DNA enzyme. All of these terminologies describe nucleic acid molecules with enzymatic activity. The specific enzymatic nucleic acid molecules described in the instant application are not limiting in the invention and those skilled in the art will recognize that all that is important in an enzymatic nucleic acid molecule of this invention is that it have a specific substrate binding site which is complementary to one or more of the target nucleic acid regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart a nucleic acid cleaving activity to the molecule (Cech et al., U.S. Patent No. 4,987,071; Cech et al., 1988, JAMA 260:20 3030-4).

By "nucleic acid molecule" as used herein is meant a molecule comprising nucleotides. The nucleic acid can be single, double, or multiple stranded and can comprise modified or unmodified nucleotides or non-nucleotides or various mixtures and combinations thereof.

By "enzymatic portion" or "catalytic domain" is meant that portion/region of the enzymatic nucleic acid molecule essential for cleavage of a nucleic acid substrate (for example see **Figures 1-5**).

By "substrate binding arm" or "substrate binding domain" is meant that portion/region of a ribozyme which is complementary to (*i.e.*, able to base-pair with) a portion of its substrate. Generally, such complementarity is 100%, but can be less if desired. For example, as few as 10 bases out of 14 may be base-paired (see for example Werner and Uhlenbeck,

1995, *Nucleic Acids Research*, 23, 2092-2096; Hammann *et al.*, 1999, *Antisense and Nucleic Acid Drug Dev.*, 9, 25-31). Such arms are shown generally in **Figures 1-5**. That is, these arms contain sequences within a ribozyme which are intended to bring ribozyme and target RNA together through complementary base-pairing interactions. The ribozyme of the invention can have binding arms that are contiguous or non-contiguous and may be of varying lengths. The length of the binding arm(s) are preferably greater than or equal to four nucleotides and of sufficient length to stably interact with the target RNA; specifically 12-100 nucleotides; more specifically 14-24 nucleotides long (see for example Werner and Uhlenbeck, *supra;* Hamman *et al., supra;* Hampel *et al.,* EP0360257; Berzal-Herrance *et al.,* 1993, *EMBO J.,* 12, 2567-73). If two binding arms are chosen, the design is such that the length of the binding arms are symmetrical (*i.e.,* each of the binding arms is of the same length; *e.g.,* five and five nucleotides, six and six nucleotides or seven and seven nucleotides long) or asymmetrical (*i.e.,* the binding arms are of different length; *e.g.,* six and three nucleotides; three and six nucleotides long; four and five nucleotides long; four and six nucleotides long; four and seven nucleotides long; and the like).

By "nuclease activating compound" is meant a compound, for example a compound having Formula I, that activates the cleavage of an RNA by a nuclease. The nuclease can comprise RNAse L. By "nuclease activating chimera" or "chimera" is meant a nuclease activating compound, for example a compound having Formula I, that is attached to a nulceic acid molecule, for example a nucleic acid molecule that binds preferentially to a target RNA. These chimeric nucleic acid molecules can comprise a nuclease activating compound and an antisense nucleic acid molecule, for example a 2',5'-oligoadenylate antisense chimera, or an enzymatic nucleic acid moleucle, for example a 2',5'-oligoadenylate enzymatic nucleic acid chimera.

By "Inozyme" or "NCH" motif or configuration is meant, an enzymatic nucleic acid molecule comprising a motif as is generally described as NCH Rz in Ludwig *et al.,* International PCT Publication No. WO 98/58058 and US Patent Application Serial No. 08/878,640. Inozymes possess endonuclease activity to cleave RNA substrates having a cleavage triplet NCH/, where N is a nucleotide, C is cytidine and H is adenosine, uridine or cytidine, and / represents the cleavage site. Inozymes can also possess endonuclease activity to cleave RNA substrates having a cleavage triplet NCN/, where N is a nucleotide, C is cytidine, and / represents the cleavage site.

By "G-cleaver" motif or configuration is meant, an enzymatic nucleic acid molecule comprising a motif as is generally described in Eckstein *et al.,* US 6,127,173 and in Kore *et al.,* 1998, *Nucleic Acids Research* 26, 4116-4120. G-cleavers possess endonuclease activity

to cleave RNA substrates having a cleavage triplet NYN/, where N is a nucleotide, Y is uridine or cytidine and / represents the cleavage site. G-cleavers can be chemically modified.

By "zinzyme" motif or configuration is meant, an enzymatic nucleic acid molecule comprising a motif as is generally described in Beigelman *et al.,* International PCT publication No. WO 99/55857 and US Patent Application Serial No. 09/918,728. Zinzymes possess endonuclease activity to cleave RNA substrates having a cleavage triplet including but not limited to, YG/Y, where Y is uridine or cytidine, and G is guanosine and / represents the cleavage site. Zinzymes can be chemically modified to increase nuclease stability through various substitutions, including substituting 2'-O-methyl guanosine nucleotides for guanosine nucleotides. In addition, differing nucleotide and/or non-nucleotide linkers can be used to substitute the 5'-gaaa-2' loop of the motif. Zinzymes represent a non-limiting example of an enzymatic nucleic acid molecule that does not require a ribonucleotide (2'-OH) group within its own nucleic acid sequence for activity.

By "amberzyme" motif or configuration is meant, an enzymatic nucleic acid molecule comprising a motif as is generally described in Beigelman *et al.,* International PCT publication No. WO 99/55857 and US Patent Application Serial No. 09/476,387. Amberzymes possess endonuclease activity to cleave RNA substrates having a cleavage triplet NG/N, where N is a nucleotide, G is guanosine, and / represents the cleavage site. Amberzymes can be chemically modified to increase nuclease stability. In addition, differing nucleoside and/or non-nucleoside linkers can be used to substitute the 5'-gaaa-3' loops of the motif. Amberzymes represent a non-limiting example of an enzymatic nucleic acid molecule that does not require a ribonucleotide (2'-OH) group within its own nucleic acid sequence for activity.

By 'DNAzyme' is meant, an enzymatic nucleic acid molecule that does not require the presence of a 2'-OH group within its own nucleic acid sequence for activity. In particular embodiments, the enzymatic nucleic acid molecule can have an attached linker or linkers or other attached or associated groups, moieties, or chains containing one or more nucleotides with 2'-OH groups. DNAzymes can be synthesized chemically or expressed endogenously *in vivo,* by means of a single stranded DNA vector or equivalent thereof. Non-limiting examples of DNAzymes are generally reviewed in Usman *et al.,* US patent No., 6,159,714; Chartrand *et al.,* 1995, *NAR* 23, 4092; Breaker *et al.,* 1995, *Chem. Bio.* 2, 655; Santoro *et al.,* 1997, *PNAS* 94, 4262; Breaker, 1999, *Nature Biotechnology,* 17, 422-423; and Santoro *et. al.,* 2000, *J. Am. Chem. Soc.,* 122, 2433-39. The "10-23" DNAzyme motif is one particular type of DNAzyme that was evolved using *in vitro* selection as generally described in Joyce *et al.,* US 5,807,718 and Santoro *et al., supra.* Additional DNAzyme motifs can be selected for

using techniques similar to those described in these references, and hence, are within the scope of the present invention.

By "nucleic acid sensor molecule" or "allozyme" as used herein is meant a nucleic acid molecule comprising an enzymatic domain and a sensor domain, where the enzymatic nucleic acid domain's ability to catalyze a chemical reaction is dependent on the interaction with a target signaling molecule, such as a nucleic acid, polynucleotide, oligonucleotide, peptide, polypeptide, or protein, for example HBV RT, HBV RT primer, or HBV Enhancer I sequence. The introduction of chemical modifications, additional functional groups, and/or linkers, to the nucleic acid sensor molecule can provide enhanced catalytic activity of the nucleic acid sensor molecule, increased binding affinity of the sensor domain to a target nucleic acid, and/or improved nuclease/chemical stability of the nucleic acid sensor molecule, and are hence within the scope of the present invention (see for example Usman *et al.,* US Patent Application No. 09/877,526, George *et al.,* US Patent Nos. 5,834,186 and 5,741,679, Shih *et al.,* US Patent No. 5,589,332, Nathan *et al.,* US Patent No 5,871,914, Nathan and Ellington, International PCT publication No. WO 00/24931, Breaker *et al.,* International PCT Publication Nos. WO 00/26226 and 98/27104, and Sullenger *et al.,* US Patent Application Serial No. 09/205,520).

By "sensor component" or "sensor domain" of the nucleic acid sensor molecule as used herein is meant, a nucleic acid sequence (e.g., RNA or DNA or analogs thereof) which interacts with a target signaling molecule, for example a nucleic acid sequence in one or more regions of a target nucleic acid molecule or more than one target nucleic acid molecule, and which interaction causes the enzymatic nucleic acid component of the nucleic acid sensor molecule to either catalyze a reaction or stop catalyzing a reaction. In the presence of target signaling molecule of the invention, such as HBV RT, HBV RT primer, or HBV Enhancer I sequence, the ability of the sensor component, for example, to modulate the catalytic activity of the nucleic acid sensor molecule, is altered or diminished in a manner that can be detected or measured. The sensor component can comprise recognition properties relating to chemical or physical signals capable of modulating the nucleic acid sensor molecule via chemical or physical changes to the structure of the nucleic acid sensor molecule. The sensor component can be derived from a naturally occurring nucleic acid binding sequence, for example, RNAs that bind to other nucleic acid sequences *in vivo*. Alternately, the sensor component can be derived from a nucleic acid molecule (aptamer), which is evolved to bind to a nucleic acid sequence within a target nucleic acid molecule. The sensor component can be covalently linked to the nucleic acid sensor molecule, or can be non-covalently associated. A person skilled in the art will recognize that all that is required is that the sensor component is able to selectively modulate the activity of the nucleic acid sensor molecule to catalyze a reaction.

By "target molecule" or "target signaling molecule" is meant a molecule capable of interacting with a nucleic acid sensor molecule, specifically a sensor domain of a nucleic acid sensor molecule, in a manner that causes the nucleic acid sensor molecule to be active or inactive. The interaction of the signaling agent with a nucleic acid sensor molecule can result in modification of the enzymatic nucleic acid component of the nucleic acid sensor molecule via chemical, physical, topological, or conformational changes to the structure of the molecule, such that the activity of the enzymatic nucleic acid component of the nucleic acid sensor molecule is modulated, for example is activated or inactivated. Signaling agents can comprise target signaling molecules such as macromolecules, ligands, small molecules, metals and ions, nucleic acid molecules including but not limited to RNA and DNA or analogs thereof, proteins, peptides, antibodies, polysaccharides, lipids, sugars, microbial or cellular metabolites, pharmaceuticals, and organic and inorganic molecules in a purified or unpurified form, for example HBV RT or HBV RT primer.

By "sufficient length" is meant a nucleic acid molecule long enough to provide the intended function under the expected condition. For example, a nucleic acid molecule of the invention needs to be of "sufficient length" to provide stable binding to a target site under the expected binding conditions and environment. In another non-limiting example, for the binding arms of an enzymatic nucleic acid, "sufficient length" means that the binding arm sequence is long enough to provide stable binding to a target site under the expected reaction conditions and environment. The binding arms are not so long as to prevent useful turnover of the nucleic acid molecule. By "stably interact" is meant interaction of the oligonucleotides with target nucleic acid (*e.g.*, by forming hydrogen bonds with complementary nucleotides in the target under physiological conditions) that is sufficient for the intended purpose (e.g., cleavage of target RNA by an enzyme).

By "equivalent" RNA to HBV or HCV is meant to include those naturally occurring RNA molecules having homology (partial or complete) to HBV or HCV proteins or encoding for proteins with similar function as HBV or HCV in various organisms, including human, rodent, primate, rabbit, pig, protozoans, fungi, plants, and other microorganisms and parasites. The equivalent RNA sequence also includes in addition to the coding region, regions such as 5'-untranslated region, 3'-untranslated region, introns, intron-exon junction and the like.

The term "component" of HBV or HCV as used herein refers to a peptide or protein subunit expressed from a HBV or HCV gene.

By "homology" is meant the nucleotide sequence of two or more nucleic acid molecules is partially or completely identical.

By "antisense nucleic acid", it is meant a non-enzymatic nucleic acid molecule that binds to target RNA by means of RNA-RNA or RNA-DNA or RNA-PNA (protein nucleic acid; Egholm *et al.,* 1993 *Nature* 365, 566) interactions and alters the activity of the target RNA (for a review, see Stein and Cheng, 1993 *Science* 261, 1004 and Woolf *et al.,* US patent No. 5,849,902). Typically, antisense molecules are complementary to a target sequence along a single contiguous sequence of the antisense molecule. However, in certain embodiments, an antisense molecule can bind to substrate such that the substrate molecule forms a loop, and/or an antisense molecule can bind such that the antisense molecule forms a loop. Thus, the antisense molecule can be complementary to two or more non-contiguous substrate sequences or two or more non-contiguous sequence portions of an antisense molecule can be complementary to a target sequence, or both. For a review of current antisense strategies, see Schmajuk *et al.,* 1999, *J. Biol. Chem.,* 274, 21783-21789, Delihas *et al.,* 1997, *Nature,* 15, 751-753, Stein *et al.,* 1997, *Antisense N. A. Drug Dev.,* 7, 151, Crooke, 2000, *Methods Enzymol.,* 313, 3-45; Crooke, 1998, *Biotech. Genet. Eng. Rev.,* 15, 121-157, Crooke, 1997, *Ad. Pharmacol.,* 40, 1-49. Antisense molecules of the instant invention can include 2-5A antisense chimera molecules. In addition, antisense DNA can be used to target RNA by means of DNA-RNA interactions, thereby activating RNase H, which digests the target RNA in the duplex. The antisense oligonucleotides can comprise one or more RNAse H activating region that is capable of activating RNAse H cleavage of a target RNA. Antisense DNA can be synthesized chemically or expressed via the use of a single stranded DNA expression vector or equivalent thereof.

By "RNase H activating region" is meant a region (generally greater than or equal to 4-25 nucleotides in length, preferably from 5-11 nucleotides in length) of a nucleic acid molecule capable of binding to a target RNA to form a non-covalent complex that is recognized by cellular RNase H enzyme (see for example Arrow *et al.,* US 5,849,902; Arrow *et al.,* US 5,989,912). The RNase H enzyme binds to the nucleic acid molecule-target RNA complex and cleaves the target RNA sequence. The RNase H activating region comprises, for example, phosphodiester, phosphorothioate (for example, at least four of the nucleotides are phosphorothiote substitutions; more specifically, 4-11 of the nucleotides are phosphorothiote substitutions), phosphorodithioate, 5'-thiophosphate, or methylphosphonate backbone chemistry or a combination thereof. In addition to one or more backbone chemistries described above, the RNase H activating region can also comprise a variety of sugar chemistries. For example, the RNase H activating region can comprise deoxyribose, arabino, fluoroarabino or a combination thereof, nucleotide sugar chemistry. Those skilled in the art will recognize that the foregoing are non-limiting examples and that any combination

of phosphate, sugar and base chemistry of a nucleic acid that supports the activity of RNase H enzyme is within the scope of the definition of the RNase H activating region and the instant invention.

By "2-5A antisense" or "2-5A antisense chimera" is meant an antisense oligonucleotide containing a 5'-phosphorylated 2'-5'-linked adenylate residue. These chimeras bind to target RNA in a sequence-specific manner and activate a cellular 2-5A-dependent ribonuclease which, in turn, cleaves the target RNA (Torrence et al., 1993 Proc. Natl. Acad. Sci. USA 90, 1300; Silverman et al., 2000, Methods Enzymol., 313, 522-533; Player and Torrence, 1998, Pharmacol. Ther., 78, 55-113).

By "triplex nucleic acid" or "triplex oligonucleotide" it is meant a polynucleotide or oligonucleotide that can bind to a double-stranded DNA in a sequence-specific manner to form a triple-strand helix. Formation of such triple helix structure has been shown to modulate transcription of the targeted gene (Duval-Valentin *et al.*, 1992, *Proc. Natl. Acad. Sci.USA*, 89, 504). Triplex nucleic acid molecules of the invention also include steric blocker nucleic acid molecules that bind to the Enhancer I region of HBV DNA (plus strand and/or minus strand) and prevent translation of HBV genomic DNA.

The term "single stranded RNA" (ssRNA) as used herein refers to a naturally occurring or synthetic ribonucleic acid molecule comprising a linear single strand, for example a ssRNA can be a messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA) etc. of a gene.

The term "single stranded DNA" (ssDNA) as used herein refers to a naturally occurring or synthetic deoxyribonucleic acid molecule comprising a linear single strand, for example, a ssDNA can be a sense or antisense gene sequence or EST (Expressed Sequence Tag).

The term "allozyme" as used herein refers to an allosteric enzymatic nucleic acid molecule, see for example George *et al.*, US Patent Nos. 5,834,186 and 5,741,679, Shih *et al.*, US Patent No. 5,589,332, Nathan *et al.*, US Patent No 5,871,914, Nathan and Ellington, International PCT publication No. WO 00/24931, Breaker *et al.*, International PCT Publication Nos. WO 00/26226 and 98/27104, and Sullenger *et al.*, International PCT publication No. WO 99/29842.

The term "2-5A chimera" as used herein refers to an oligonucleotide containing a 5'-phosphorylated 2'-5'-linked adenylate residue. These chimeras bind to target RNA in a sequence-specific manner and activate a cellular 2-5A-dependent ribonuclease which, in turn, cleaves the target RNA (Torrence *et al.*, 1993 *Proc. Natl. Acad. Sci. USA* 90, 1300;

Silverman *et al.*, 2000, *Methods Enzymol.*, 313, 522-533; Player and Torrence, 1998, *Pharmacol. Ther.*, 78, 55-113).

The term "double stranded RNA" or "dsRNA" as used herein refers to a double stranded RNA molecule capable of RNA interference "RNAi", including short interfering RNA "siRNA" see for example Bass, 2001, *Nature*, 411, 428-429; Elbashir et al., 2001, *Nature*, 411, 494-498; and Kreutzer *et al.*, International PCT Publication No. WO 00/44895; Zernicka-Goetz *et al.*, International PCT Publication No. WO 01/36646; Fire, International PCT Publication No. WO 99/32619; Plaetinck *et al.*, International PCT Publication No. WO 00/01846; Mello and Fire, International PCT Publication No. WO 01/29058; Deschamps-Depaillette, International PCT Publication No. WO 99/07409; and Li *et al.*, International PCT Publication No. WO 00/44914.

By "gene" it is meant, a nucleic acid that encodes an RNA, for example, nucleic acid sequences including, but not limited to, structural genes encoding a polypeptide.

By "complementarity" is meant that a nucleic acid can form hydrogen bond(s) with another nucleic acid sequence by either traditional Watson-Crick or other non-traditional types. In reference to the nucleic molecules of the present invention, the binding free energy for a nucleic acid molecule with its target or complementary sequence is sufficient to allow the relevant function of the nucleic acid to proceed, e.g., ribozyme cleavage, antisense or triple helix modulation. Determination of binding free energies for nucleic acid molecules is well known in the art (see, e.g., Turner et al., 1987, *CSH Symp. Quant. Biol.* LII pp.123-133; Frier et al., 1986, *Proc. Nat. Acad. Sci.* USA 83:9373-9377; Turner et al., 1987, *J. Am. Chem. Soc.* 109:3783-3785). A percent complementarity indicates the percentage of contiguous residues in a nucleic acid molecule that can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence (e.g., 5, 6, 7, 8, 9, 10 out of 10 being 50%, 60%, 70%, 80%, 90%, and 100% complementary). "Perfectly complementary" means that all the contiguous residues of a nucleic acid sequence will hydrogen bond with the same number of contiguous residues in a second nucleic acid sequence.

As used herein "cell" is used in its usual biological sense, and does not refer to an entire multicellular organism, e.g., specifically does not refer to a human. The cell can be present in an organism, e.g., birds, plants and mammals such as humans, cows, sheep, apes, monkeys, swine, dogs, and cats. The cell can be prokaryotic (e.g., bacterial cell) or eukaryotic (e.g., mammalian or plant cell).

By "HBV proteins" or "HCV proteins" is meant, a protein or a mutant protein derivative thereof, comprising sequence expressed and/or encoded by the HBV genome.

By "highly conserved sequence region" is meant a nucleotide sequence of one or more regions in a target gene does not vary significantly from one generation to the other or from one biological system to the other.

By "highly conserved nucleic acid binding region" is meant an amino acid sequence of one or more regions in a target protein that does not vary significantly from one generation to the other or from one biological system to the other.

By "related to the levels of HBV" is meant that the reduction of HBV expression (specifically HBV gene) RNA levels and thus reduction in the level of the respective protein will relieve, to some extent, the symptoms of the disease or condition.

By "related to the levels of HCV" is meant that the reduction of HCV expression (specifically HCV gene) RNA levels and thus reduction in the level of the respective protein will relieve, to some extent, the symptoms of the disease or condition.

By "RNA" is meant a molecule comprising at least one ribonucleotide residue. By "ribonucleotide" is meant a nucleotide with a hydroxyl group at the 2' position of a β-D-ribo-furanose moiety.

By "vector" is meant any nucleic acid- and/or viral-based technique used to express and/or deliver a desired nucleic acid.

By "patient" is meant an organism, which is a donor or recipient of explanted cells or the cells themselves. "Patient" also refers to an organism to which the nucleic acid molecules of the invention can be administered. In one embodiment, a patient is a mammal or mammalian cells. In another embodiment, a patient is a human or human cells.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

First the drawings will be described briefly.

<u>Drawings</u>

Figure 1 shows the secondary structure model for seven different classes of enzymatic nucleic acid molecules. Arrow indicates the site of cleavage. -------- indicate the target sequence. Lines interspersed with dots are meant to indicate tertiary interactions. - is meant to

indicate base-paired interaction. **Group I Intron**: P1-P9.0 represent various stem-loop structures (Cech *et al.*, 1994, *Nature Struc. Bio.*, 1, 273). **RNase P (M1RNA)**: EGS represents external guide sequence (Forster *et al.*, 1990, *Science*, 249, 783; Pace *et al.*, 1990, *J. Biol. Chem.*, 265, 3587). **Group II Intron:** 5'SS means 5' splice site; 3'SS means 3'-splice site; IBS means intron binding site; EBS means exon binding site (Pyle *et al.*, 1994, *Biochemistry*, 33, 2716). **VS RNA:** I-VI are meant to indicate six stem-loop structures; shaded regions are meant to indicate tertiary interaction (Collins, International PCT Publication No. WO 96/19577). **HDV Ribozyme:** I-IV are meant to indicate four stem-loop structures (Been *et al.*, US Patent No. 5,625,047). **Hammerhead Ribozyme:** I-III are meant to indicate three stem-loop structures; stems I-III can be of any length and may be symmetrical or asymmetrical (Usman *et al.*, 1996, *Curr. Op. Struct. Bio.*, 1, 527). **Hairpin Ribozyme:** Helix 1, 4 and 5 can be of any length; Helix 2 is between 3 and 8 base-pairs long; Y is a pyrimidine; Helix 2 (H2) is provided with a least 4 base pairs (*i.e.*, n is 1, 2, 3 or 4) and helix 5 can be optionally provided of length 2 or more bases (preferably 3 - 20 bases, *i.e.*, m is from 1 - 20 or more). Helix 2 and helix 5 may be covalently linked by one or more bases (*i.e.*, r is ≥ 1 base). Helix 1, 4 or 5 may also be extended by 2 or more base pairs (*e.g.*, 4 - 20 base pairs) to stabilize the ribozyme structure, and preferably is a protein binding site. In each instance, each N and N' independently is any normal or modified base and each dash represents a potential base-pairing interaction. These nucleotides may be modified at the sugar, base or phosphate. Complete base-pairing is not required in the helices, but is preferred. Helix 1 and 4 can be of any size (*i.e.*, o and p is each independently from 0 to any number, *e.g.*, 20) as long as some base-pairing is maintained. Essential bases are shown as specific bases in the structure, but those in the art will recognize that one or more may be modified chemically (abasic, base, sugar and/or phosphate modifications) or replaced with another base without significant effect. Helix 4 can be formed from two separate molecules, *i.e.*, without a connecting loop. The connecting loop when present may be a ribonucleotide with or without modifications to its base, sugar or phosphate. "q" ≥ is 2 bases. The connecting loop can also be replaced with a non-nucleotide linker molecule. H refers to bases A, U, or C. Y refers to pyrimidine bases. "_____" refers to a covalent bond. (Burke *et al.*, 1996, *Nucleic Acids & Mol. Biol.*, 10, 129; Chowrira *et al.*, US Patent No. 5,631,359).

Figure 2 shows examples of chemically stabilized ribozyme motifs. **HH Rz**, represents hammerhead ribozyme motif (Usman *et al.*, 1996, *Curr. Op. Struct. Bio.*, 1, 527); **NCH Rz** represents the NCH ribozyme motif (Ludwig & Sproat, International PCT Publication No. WO 98/58058); **G-Cleaver**, represents G-cleaver ribozyme motif (Kore *et al.*, 1998, *Nucleic Acids Research*, 26, 4116-4120). N or **n**, represent independently a nucleotide which may be same or different and have complementarity to each other; **rI**, represents ribo-Inosine nucleotide; arrow indicates the site of cleavage within the target. Position 4 of the HH Rz and the NCH Rz is shown as having 2'-*C*-allyl modification, but

those skilled in the art will recognize that this position can be modified with other modifications well known in the art, so long as such modifications do not significantly inhibit the activity of the ribozyme.

Figure 3 shows an example of the Amberzyme ribozyme motif that is chemically stabilized (see, for example, Beigelman *et al.,* International PCT publication No. WO 99/55857; also referred to as Class I Motif). The Amberzyme motif is a class of enzymatic nucleic acid molecules that do not require the presence of a ribonucleotide (2'-OH) group for activity.

Figure 4 shows an example of the Zinzyme A ribozyme motif that is chemically stabilized (see, for example, International PCT publication No. WO 99/55857; also referred to as Class A Motif). The Zinzyme motif is a class of enzymatic nucleic acid molecules that do not require the presence of a ribonucleotide (2'-OH) group for activity.

Figure 5 shows an example of a DNAzyme motif described by Santoro *et al.,* 1997, *PNAS,* 94, 4262.

Figure 6 is a bar graph showing the percent change in serum HBV DNA levels following fourteen days of ribozyme treatment in HBV transgenic mice. Ribozymes targeting sites 273 (RPI.18341) and 1833 (RPI.18371) of HBV RNA administerd via continuous s.c. infusion at 10, 30, and 100 mg/kg/day are compared to continuous s.c. infusion administration of scrambled attenuated core ribozyme and saline controls, and orally administered 3TC® (300 mg/kg/day) and saline controls.

Figure 7 is a bar graph showing the mean serum HBV DNA levels following fourteen days of ribozyme treatment in HBV transgenic mice. Ribozymes targeting sites 273 (RPI.18341) and 1833 (RPI.18371) of HBV RNA administerd via continuous s.c. infusion at 10, 30, and 100 mg/kg/day are compared to continuous s.c. infusion administration of scrambled attenuated core ribozyme and saline controls, and orally administered 3TC® (300 mg/kg/day) and saline controls.

Figure 8 is a bar graph showing the decrease in serum HBV DNA (log) levels following fourteen days of ribozyme treatment in HBV transgenic mice. Ribozymes targeting sites 273 (RPI.18341) and 1833 (RPI.18371) of HBV RNA administerd via continuous s.c. infusion at 10, 30, and 100 mg/kg/day are compared to continuous s.c. infusion administration of scrambled attenuated core ribozyme and saline controls, and orally administered 3TC® (300 mg/kg/day) and saline controls.

Figure 9 is a bar graph showing the decrease in HBV DNA in HepG2.2.15 cells after treatment with ribozymes targeting sites 273 (RPI.18341), 1833 (RPI.18371), 1874

(RPI.18372), and 1873 (RPI.18418) of HBV RNA as compared to a scrambled attenuated core ribozyme (RPI.20995).

Figure 10 is a bar graph showing reduction in HBsAg levels following treatment of HepG2 cells with anti-HBV arm, stem, and loop-variant ribozymes (RPI.18341, RPI.22644, RPI.22645, RPI.22646, RPI.22647, RPI.22648, RPI.22649, and RPI.22650) targeting site 273 of the HBV pregenomic RNA as compared to a scrambled attenuated core ribozyme (RPI.20599).

Figure 11 is a bar graph showing reduction in HBsAg levels following treatment of HepG2 cells with RPI 18341 alone or in combination with Infergen®. At either 500 or 1000 units of Infergen®, the addition of 200 nM of RPI.18341 results in a 75-77% increase in anti-HBV activity as judged by the level of HBsAg secreted from the treated Hep G2 cells. Conversely, the anti-HBV activity of RPI.18341(at 200 nM) is increased 31-39% when used in combination of 500 or 1000 units of Infergen®.

Figure 12 is a bar graph showing reduction in HBsAg levels following treatment of HepG2 cells with RPI 18341 alone or in combination with Lamivudine. At 25 nM Lamivudine (3TC®), the addition of 100 nM of RPI.18341 results in a 48% increase in anti-HBV activity as judged by the level of HBsAg secreted from treated Hep G2 cells. Conversely, the anti-HBV activity of RPI.18341 (at 100 nM) is increased 31% when used in combination with 25 nM Lamivudine.

Figure 13 shows a scheme which outlines the steps involved in HBV reverse transcription. The HBV polymerase/reverse transcriptase binds to the 5'-stem-loop of the HBV pregenomic RNA and synthesizes a primer from the UUCA template. The reverse transcriptase and tetramer primer are translocated to the 3'-DR1 site. The RT primer binds to the UUCA sequence in the DR1 element and minus strand synthesis begins.

Figure 14 shows a non-limiting example of inhibition of HBV reverse transcription. A decoy molecule binds to the HBV RT primer, thereby preventing translocation of the RT to the 3'-DR1 site and preventing minus strand synthesis.

Figure 15 shows data of a HBV nucleic acid screen of 2'-O-allyl modified nucleic acid molecules. The levels of HbsAg were determined by ELISA. Inhibition of HBV is correlated to HBsAg antigen levels.

Figure 16 shows data of a HBV nucleic acid screen of 2'-O-methyl modified nucleic acid molecules. The levels of HbsAg were determined by ELISA. Inhibition of HBV is correlated to HBsAg antigen levels.

Figure 17 shows dose response data of 2'-O-methyl modified nucleic acid molecules targeting the HBV reverse transcriptase primer compared to levels of HBsAg.

Figure 18 shows data of nucleic acid screen of nucleic acid molecules (200 nM) targeting the HBV Enhancer I core region compared to levels of HBsAg.

Figure 19 shows data of nucleic acid screen of nucleic acid molecules (400 nM) targeting the HBV Enhancer I core region compared to levels of HBsAg.

Figure 20 shows dose response data of nucleic acid molecules targeting the HBV Enhancer I core region compared to levels of HBsAg.

Figure 21 shows a graph depicting HepG2.2.15 tumor growth in athymic nu/nu female mice as tumor volume (mm^3) vs time (days).

Figure 22 shows a graph depicting HepG2.2.15 tumor growth in athymic nu/nu female mice as tumor volume (mm^3) vs time (days). Inoculated HepG2.2.15 cells were selected for antibiotic resistance to G418 before introduction into the mouse.

Figure 23 is a schematic representation of the Dual Reporter System utilized to demonstrate enzymatic nucleic acid mediated reduction of luciferase activity in cell culture.

Figure 24 shows a schematic view of the secondary structure of the HCV 5'UTR (Brown *et al.*, 1992, *Nucleic Acids Res.*, 20, 5041-45; Honda *et al.*, 1999, *J. Virol.*, 73, 1165-74). Major structural domains are indicated in bold. Enzymatic nucleic acid cleavage sites are indicated by arrows. Solid arrows denote sites amenable to amino-modified enzymatic nucleic acid inhibition. Lead cleavage sites (195 and 330) are indicated with oversized solid arrows.

Figure 25 shows a non-limiting example of a nuclease resistant enzymatic nucleic acid molecule. Binding arms are indicated as stem I and stem III. Nucleotide modifications are indicated as follows: 2'-O-methyl nucleotides, lowercase; ribonucleotides, uppercase G, A; 2'-amino-uridine, u; inverted 3'-3' deoxyabasic, **B**. The positions of phosphorothioate linkages at the 5'-end of each enzymatic nucleic acid are indicated by subscript "s". *H* indicates A, C or U ribonucleotide, *N'* indicates A, C G or U ribonucleotide in substrate, *n* indicates base complementary to the N'. The U4 and U7 positions in the catalytic core are indicated.

Figure 26 is a set of bar graphs showing enzymatic nucleic acid mediated inhibition of HCV-luciferase expression in OST7 cells. OST7 cells were transfected with complexes containing reporter plasmids (2 µg/mL), enzymatic nucleic acids (100 nM) and lipid. The ratio of HCV-firefly luciferase luminescence/Renilla luciferase luminescence was determined

for each enzymatic nucleic acid tested and was compared to treatment with the ICR, an irrelevant control enzymatic nucleic acid lacking specificity to the HCV 5'UTR (adjusted to 1). Results are reported as the mean of triplicate samples $\pm$ SD. In **Figure 26A**, OST7 cells were treated with enzymatic nucleic acids (100 nM) targeting conserved sites (indicated by cleavage site) within the HCV 5'UTR. In **Figure 26B**, OST7 cells were treated with a subset of enzymatic nucleic acids to lead HCV sites (indicated by cleavage site) and corresponding attenuated core (AC) controls. Percent decrease in firefly/Renilla luciferase ratio after treatment with active enzymatic nucleic acids as compared to treatment with corresponding ACs is shown when the decrease is $\geq$ 50% and statistically significant. Similar results were obtained with 50 nM enzymatic nucleic acid.

Figure 27 is a series of line graphs showing the dose-dependent inhibition of HCV/luciferase expression following enzymatic nucleic acid treatment. Active enzymatic nucleic acid was mixed with corresponding AC to maintain a 100 nM total oligonucleotide concentration and the same lipid charge ratio. The concentration of active enzymatic nucleic acid for each point is shown. **Figure 27A–E** shows enzymatic nucleic acids targeting sites 79, 81, 142, 195, or 330, respectively. Results are reported as the mean of triplicate samples $\pm$ SD.

Figure 28 is a set of bar graphs showing reduction of HCV/luciferase RNA and inhibition of HCV-luciferase expression in OST7 cells. OST7 cells were transfected with complexes containing reporter plasmids (2 μg /ml), enzymatic nucleic acids, BACs or SACs (50 nM) and lipid. Results are reported as the mean of triplicate samples $\pm$ SD. In **Figure 28A** the ratio of HCV-firefly luciferase RNA/Renilla luciferase RNA is shown for each enzymatic nucleic acid or control tested. As compared to paired BAC controls (adjusted to 1), luciferase RNA levels were reduced by 40% and 25% for the site 195 or 330 enzymatic nucleic acids, respectively. In **Figure 28B** the ratio of HCV-firefly luciferase luminescence/Renilla luciferase luminescence is shown after treatment with site 195 or 330 enzymatic nucleic acids or paired controls. As compared to paired BAC controls (adjusted to 1), inhibition of protein expression was 70% and 40% for the site 195 or 330 enzymatic nucleic acids, respectively $P < 0.01$.

Figure 29 is a set a bar graphs showing interferon (IFN) alpha 2a and 2b dose response in combination with site 195 anti-HCV enzymatic nucleic acid treatment. **Figure 29A** shows data for IFN alfa 2a treatment. **Figure 29B** shows data for IFN alfa 2b treatment. Viral yield is reported from HeLa cells pretreated with IFN in units/ml (U/ml) as indicated for 4 h prior to infection and then treated with either 200 nM control (SAC) or site 195 anti-HCV enzymatic nucleic acid (195 RZ) for 24 h after infection. Cells were infected with a MOI =

0.1 for 30 min and collected at 24 h post infection. Error bars represent the S.D. of the mean of triplicate determinations.

Figure 30 is a line graph showing site 195 anti-HCV enzymatic nucleic acid dose response in combination with interferon (IFN) alpha 2a and 2b pretreatment. Viral yield is reported from HeLa cells pretreated for 4 h with or without IFN and treated with doses of site 195 anti-HCV enzymatic nucleic acid (195 RZ) as indicated for 24 h after infection. Anti-HCV enzymatic nucleic acid was mixed with control oligonucleotide (SAC) to maintain a constant 200 nM total dose of nucleic acid for delivery. Cells were infected with a MOI = 0.1 for 30 min and collected at 24 h post infection. Error bars represent the S.D. of the mean of triplicate determinations.

Figure 31 is a set of bar graphs showing data from consensus interferon (CIFN)/enzymatic nucleic acid combination treatment. **Figure 31A** shows CIFN dose response with site 195 anti-HCV enzymatic nucleic acid treatment. Viral yield is reported from cells pretreated with CIFN in units/ml (U/ml) as indicated and treated with either 200 nM control (SAC) or site 195 anti-HCV enzymatic nucleic acid (195 RZ). **Figure 31B** shows site 195 anti-HCV enzymatic nucleic acid dose response with CIFN pretreatment. Viral yield is reported from cells pretreated with or without CIFN and treated with concentrations of site 195 anti-HCV enzymatic nucleic acid (195 RZ) as indicated. Anti-HCV enzymatic nucleic acid was mixed with control oligonucleotide (SAC) to maintain a constant 200 nM total dose of nucleic acid for delivery. Cells were infected with a MOI = 0.1 for 30 min. and collected at 24 h post infection. Error bars represent the S.D. of the mean of triplicate determinations.

Figure 32 is a bar graph showing enzymatic nucleic acid activity and enhanced antiviral effect of an anti-HCV enzymatic nucleic acid targeting site 195 used in combination with consensus interferon (CIFN). Viral yield is reported from cells treated as indicated. BAC, cells were treated with 200 nM BAC (binding attenuated control) for 24 h after infection; CIFN+BAC, cells were treated with 12.5 U/ml CIFN for 4 h prior to infection and with 200 nM BAC for 24 h after infection; 195 RZ, cells were treated with 200 nM site 195 anti-HCV enzymatic nucleic acid for 24 h after infection; CIFN + 195 RZ, cells were treated with 12.5 U/ml CIFN for 4 h prior to infection and with 200 nM site 195 anti-HCV enzymatic nucleic acid for 24 h after infection. Cells were infected with a MOI = 0.1 for 30 min. Error bars represent the S.D. of the mean of triplicate determinations.

Figure 33 is a bar graph showing inhibition of a HCV-PV chimera replication by treatment with zinzyme enzymatic nucleic acid molecules targeting different sites within the HCV 5'-UTR compared to a scrambled attenuated core control (SAC) zinzyme.

Figure 34 is a bar graph showing inhibition of a HCV-PV chimera replication by antisense nucleic acid molecules targeting conserved regions of the HCV 5'-UTR compared to scrambled antisense controls.

Figure 35 shows the structure of compounds (2-5A) utilized in the study. "X" denotes the position of oxygen (O) in analog I or sulfur (S) in thiophosphate (P=S) analog II. The 2-5A compounds were synthesized, deprotected and purified as described herein utilizing CPG support with 3'-inverted abasic nucleotide. For chain extension 5'-O-(4,4'-dimetoxytrityl)-3'-O-(tert-butyldimethylsilyl)-N6-benzoyladenosine-2-cyanoethyl-N,N-diisopropyl-phosphoramidite (Chem. Genes Corp., Waltham, MA) was employed. Introduction of a 5'-terminal phosphate (analog I) or thiophosphate (analog II) group was performed with "Chemical Phosphorylation Reagent" (Glen Research, Sterling , VA). Structures of the final compounds were confirmed by MALDI-TOF analysis.

Figure 36 is a bar graph showing ribozyme activity and enhanced antiviral effect. (A) Interferon/ribozyme combination treatment. (B) 2-5A/ribozyme combination treatment. HeLa cells seeded in 96-well plates (10,000 cells per well) were pretreated as indicated for 4 hours. For pretreatment, SAC (RPI 17894), RZ (RPI 13919), and 2-5A analog I (RPI 21096) (200 nM) were complexed with lipid cytofectin. Cells were then infected with HCV-PV at a multiplicity of infection of 0.1. Virus inoculum was replaced after 30 minutes with media containing 5% serum and 100 nM RZ or SAC as indicated, complexed with cytofectin RPI.9778. After 20 hours, cells were lysed by 3 freeze/thaw cycles and virus was quantified by plaque assay. Plaque forming units (PFU)/ml are shown as the mean of triplicate samples + SEM. The absolute amount of viral yield in treated cells varied from day to day, presumably due to day to day variations in cell plating and transfection complexation. None, normal media; IFN, 10 U/ml consensus interferon; SAC, scrambled arm attenuated core control (RPI 17894); RZ, anti-HCV ribozyme (RPI 13919); 2-5A, (RPI 21096).

Figure 37 is a graph showing the inhibition of viral replication with anti-HCV ribozyme (RPI 13919) or 2-5A (RPI 21096) treatment. HeLa cells were treated as described in **Figure 36** except that there was no pretreatment and 200 nM oligonucleotide was used for treatment. 2-5A P=S contains a 5'-terminal thiophosphate (RPI21095) (see **Figure 35**).

Figure 38 is a bar graph showing anti-HCV ribozyme in combination with 2-5A treatment. HeLa cells were treated as described in **Figure 37** except concentrations were co-varied as shown to maintain a constant 200 nM total oligonucleotide dose for transfection. Cells treated with 50 nM anti-HCV ribozyme (RPI 13919) (middle bars) were also treated with 150 nM SAC (RPI 17894) or 2-5A (RPI 21096); likewise, cells treated with 100 nM anti-HCV ribozyme (bars at right) were also treated with 100 nM SAC or 2-5A.

Mechanism of action of Nucleic Acid Molecules of the Invention

Decoy: Nucleic acid decoy molecules are mimetics of naturally occurring nucleic acid molecules or portions of naturally occurring nucleic acid molecules that can be used to modulate the function of a specific protein or a nucleic acid whose activity is dependant on interaction with the naturally occurring nucleic acid molecule. Decoys modulate the function of a target protein or nucleic acid by competing with authentic nucleic acid binding to the ligand of interest. Often, the nucleic acid decoy is a truncated version of a nucleic acid sequence that is recognized, for example by a particular protein, such as a transcription factor or polymerase. Decoys can be chemically modified to increase binding affinity to the target ligand as well as to increase the enzymatic and chemical stability of the decoy. In addition, bridging and non-bridging linkers can be introduced into the decoy sequence to provide additional binding affinity to the target ligand. Decoy molecules of the invention that bind to an HCV or HBV target, such as HBV reverse transcriptase or HBV reverse transcriptase primer, or an enhancer region of the HBV pregenomic RNA, for example the Enhancer I element, modulate the transcription of RNA to DNA and therefore modulate expression of the pregenomic RNA of the virus (see **Figures 13 and 14**).

Aptamer: Nucleic acid aptamers can be selected to specifically bind to a particular ligand of interest (see for example Gold *et al.*, US 5,567,588 and US 5,475,096, Gold *et al.*, 1995, *Annu. Rev. Biochem.*, 64, 763; Brody and Gold, 2000, *J. Biotechnol.*, 74, 5; Sun, 2000, *Curr. Opin. Mol. Ther.*, 2, 100; Kusser, 2000, *J. Biotechnol.*, 74, 27; Hermann and Patel, 2000, *Science*, 287, 820; and Jayasena, 1999, *Clinical Chemistry*, 45, 1628). For example, the use of in vitro selection can be applied to evolve nucleic acid aptamers with binding specificity for HBV RT and/or HBV RT primer. Nucleic acid aptamers can include chemical modifications and linkers as described herein. Aptamer molecules of the invention that bind to a reverse transcriptase or reverse transcriptase primer, such as HBV reverse transcriptase or HBV reverse transcriptase primer, modulate the transcription of RNA to DNA and therefore modulate expression of the pregenomic RNA of the virus.

Antisense: Antisense molecules can be modified or unmodified RNA, DNA, or mixed polymer oligonucleotides and primarily function by specifically binding to matching sequences resulting in modulation of peptide synthesis (Wu-Pong, Nov 1994, *BioPharm*, 20-33). The antisense oligonucleotide binds to target RNA by Watson Crick base-pairing and blocks gene expression by preventing ribosomal translation of the bound sequences either by steric blocking or by activating RNase H enzyme. Antisense molecules can also alter protein synthesis by interfering with RNA processing or transport from the nucleus into the cytoplasm (Mukhopadhyay & Roth, 1996, *Crit. Rev. in Oncogenesis* 7, 151-190).

In addition, binding of single stranded DNA to RNA may result in nuclease degradation of the heteroduplex (Wu-Pong, *supra*; Crooke, *supra*). To date, the only backbone modified DNA chemistry which will act as substrates for RNase H are phosphorothioates, phosphorodithioates, and borontrifluoridates. Recently, it has been reported that 2'-arabino and 2'-fluoro arabino- containing oligos can also activate RNase H activity.

A number of antisense molecules have been described that utilize novel configurations of chemically modified nucleotides, secondary structure, and/or RNase H substrate domains (Woolf *et al.*, International PCT Publication No. WO 98/13526; Thompson *et al.*, USSN 60/082,404 which was filed on April 20, 1998; Hartmann *et al.*, USSN 60/101,174 which was filed on September 21, 1998) all of these are incorporated by reference herein in their entirety.

Antisense DNA can be used to target RNA by means of DNA-RNA interactions, thereby activating RNase H, which digests the target RNA in the duplex. Antisense DNA can be chemically synthesized or can be expressed via the use of a single stranded DNA intracellular expression vector or the equivalent thereof.

Triplex Forming Oligonucleotides (TFO): Single stranded oligonucleotide can be designed to bind to genomic DNA in a sequence specific manner. TFOs can be comprised of pyrimidine-rich oligonucleotides which bind DNA helices through Hoogsteen Base-pairing (Wu-Pong, *supra*). In addition, TFOs can be chemically modified to increase binding affinity to target DNA sequences. The resulting triple helix composed of the DNA sense, DNA antisense, and TFO disrupts RNA synthesis by RNA polymerase. The TFO mechanism can result in gene expression or cell death since binding may be irreversible (Mukhopadhyay & Roth, *supra*)

2'-5' Oligoadenylates: The 2-5A system is an interferon-mediated mechanism for RNA degradation found in higher vertebrates (Mitra *et al.*, 1996, *Proc Nat Acad Sci USA* 93, 6780-6785). Two types of enzymes, 2-5A synthetase and RNase L, are required for RNA cleavage. The 2-5A synthetases require double stranded RNA to form 2'-5' oligoadenylates (2-5A). 2-5A then acts as an allosteric effector for utilizing RNase L, which has the ability to cleave single stranded RNA. The ability to form 2-5A structures with double stranded RNA makes this system particularly useful for modulation of viral replication.

(2'-5') oligoadenylate structures can be covalently linked to antisense molecules to form chimeric oligonucleotides capable of RNA cleavage (Torrence, *supra*). These molecules putatively bind and activate a 2-5A-dependent RNase, the oligonucleotide/enzyme complex then binds to a target RNA molecule which can then be cleaved by the RNase enzyme. The covalent attachment of 2'-5' oligoadenylate structures is not limited to

antisense applications, and can be further elaborated to include attachment to nucleic acid molecules of the instant invention.

RNA interference (RNAi): RNA interference refers to the process of sequence specific post transcriptional gene silencing in animals mediated by short interfering RNAs (siRNA) (Fire *et al.*, 1998, *Nature*, 391, 806). The corresponding process in plants is commonly referred to as post transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi. The process of post transcriptional gene silencing is thought to be an evolutionarily conserved cellular defense mechanism used to prevent the expression of foreign genes which is commonly shared by diverse flora and phyla (Fire *et al.*, 1999, *Trends Genet.*, 15, 358). Such protection from foreign gene expression may have evolved in response to the production of double stranded RNAs (dsRNA) derived from viral infection or the random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single stranded RNA or viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response though a mechanism that has yet to be fully characterized. This mechanism appears to be different from the interferon response that results from dsRNA mediated activation of protein kinase PKR and 2',5'-oligoadenylate synthetase resulting in non-specific cleavage of mRNA by ribonuclease L.

The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNA) (Berstein *et al.*, 2001, *Nature*, 409, 363). Short interfering RNAs derived from dicer activity are typically about 21-23 nucleotides in length and comprise about 19 base pair duplexes. Dicer has also been implicated in the excision of 21 and 22 nucleotide small temporal RNAs (stRNA) from precursor RNA of conserved structure that are implicated in translational control (Hutvagner *et al.*, 2001, *Science*, 293, 834). The RNAi response also features an endonuclease complex containing a siRNA, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single stranded RNA having sequence homologous to the siRNA. Cleavage of the target RNA takes place in the middle of the region complementary to the guide sequence of the siRNA duplex (Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188).

Short interfering RNA mediated RNAi has been studied in a variety of systems. Fire *et al.*, 1998, *Nature*, 391, 806, were the first to observe RNAi in *C. Elegans*. Wianny and Goetz, 1999, *Nature Cell Biol.*, 2, 70, describes RNAi mediated by dsRNA in mouse embryos. Hammond *et al.*, 2000, *Nature*, 404, 293, describe RNAi in *Drosophila* cells transfected with dsRNA. Elbashir *et al.*, 2001, *Nature*, 411, 494, describe RNAi induced by introduction of duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells. Recent work in Drosophila embryonic lysates has revealed certain requirements for siRNA length, structure, chemical composition,

and sequence that are essential to mediate efficient RNAi activity. These studies have shown that 21 nucleotide siRNA duplexes are most active when containing two nucleotide 3'-overhangs. Furthermore, substitution of one or both siRNA strands with 2'-deoxy or 2'-O-methyl nucleotides abolishes RNAi activity, whereas substitution of 3'-terminal siRNA nucleotides with deoxy nucleotides was shown to be tolerated. Mismatch sequences in the center of the siRNA duplex were also shown to abolish RNAi activity. In addition, these studies also indicate that the position of the cleavage site in the target RNA is defined by the 5'-end of the siRNA guide sequence rather than the 3'-end (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877). Other studies have indicated that a 5'-phosphate on the target-complementary strand of a siRNA duplex is required for siRNA activity and that ATP is utilized to maintain the 5'-phosphate moiety on the siRNA (Nykanen *et al.*, 2001, *Cell,* 107, 309), however siRNA molecules lacking a 5'-phosphate are active when introduced exogenously, suggesting that 5'-phosphorylation of siRNA constructs may occur *in vivo*.

Enzymatic Nucleic Acid: Several varieties of naturally occurring enzymatic RNAs are presently known (Doherty and Doudna, 2001, *Annu. Rev. Biophys. Biomol. Struct.*, 30, 457-475; Symons, 1994, *Curr. Opin. Struct. Biol.*, 4, 322-30). In addition, several *in vitro* selection (evolution) strategies (Orgel, 1979, *Proc. R. Soc. London,* B 205, 435) have been used to evolve new nucleic acid catalysts capable of catalyzing cleavage and ligation of phosphodiester linkages (Joyce, 1989, *Gene,* 82, 83-87; Beaudry *et al.*, 1992, *Science* 257, 635-641; Joyce, 1992, *Scientific American* 267, 90-97; Breaker *et al.*, 1994, *TIBTECH* 12, 268; Bartel *et al.*, 1993, *Science* 261:1411-1418; Szostak, 1993, *TIBS* 17, 89-93; Kumar *et al.*, 1995, *FASEB J.*, 9, 1183; Breaker, 1996, *Curr. Op. Biotech.*, 7, 442; Santoro *et al.*, 1997, *Proc. Natl. Acad. Sci.*, 94, 4262; Tang *et al.*, 1997, *RNA* 3, 914; Nakamaye & Eckstein, 1994, *supra*; Long & Uhlenbeck, 1994, supra; Ishizaka et al., 1995, *supra*; Vaish *et al.*, 1997, *Biochemistry* 36, 6495). Each can catalyze a series of reactions including the hydrolysis of phosphodiester bonds in *trans* (and thus can cleave other RNA molecules) under physiological conditions.

Nucleic acid molecules of this invention can block HBV or HCV protein expression and can be used to treat disease or diagnose disease associated with the levels of HBV or HCV.

The enzymatic nature of an enzymatic nucleic acid has significant advantages, such as the concentration of nucleic acid necessary to affect a therapeutic treatment is low. This advantage reflects the ability of the enzymatic nucleic acid molecule to act enzymatically. Thus, a single enzymatic nucleic acid molecule is able to cleave many molecules of target RNA. In addition, the enzymatic nucleic acid molecule is a highly specific modulator, with the specificity of modulation depending not only on the base-pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches,

or base-substitutions, near the site of cleavage can be chosen to completely eliminate catalytic activity of an enzymatic nucleic acid molecule.

Nucleic acid molecules having an endonuclease enzymatic activity are able to repeatedly cleave other separate RNA molecules in a nucleotide base sequence-specific manner. With proper design and construction, such enzymatic nucleic acid molecules can be targeted to any RNA transcript, and efficient cleavage achieved *in vitro* (Zaug *et al.*, 324, *Nature* 429 1986; Uhlenbeck, 1987 *Nature* 328, 596; Kim et al., 84 *Proc. Natl. Acad. Sci. USA* 8788, 1987; Dreyfus, 1988, *Einstein Quart. J. Bio. Med.*, 6, 92; Haseloff and Gerlach, 334 *Nature* 585, 1988; Cech, 260 *JAMA* 3030, 1988; and Jefferies et al., 17 *Nucleic Acids Research* 1371, 1989; Chartrand *et al.*, 1995, *Nucleic Acids Research* 23, 4092; Santoro *et al.*, 1997, *PNAS* 94, 4262).

Because of their sequence specificity, *trans*-cleaving enzymatic nucleic acid molecules show promise as therapeutic agents for human disease (Usman & McSwiggen, 1995 *Ann. Rep. Med. Chem.* 30, 285-294; Christoffersen and Marr, 1995 *J. Med. Chem.* 38, 2023-2037). Enzymatic nucleic acid molecule can be designed to cleave specific RNA targets within the background of cellular RNA. Such a cleavage event renders the RNA non-functional and abrogates protein expression from that RNA. In this manner, synthesis of a protein associated with a disease state can be selectively modulated(Warashina *et al.,* 1999, *Chemistry and Biology,* 6, 237-250.

The present invention also features nucleic acid sensor molecules or allozymes having sensor domains comprising nucleic acid decoys and/or aptamers of the invention. Interaction of the nucleic acid sensor molecule's sensor domain with a molecular target, such as HCV or HBV target, e.g., HBV RT and/or HBV RT primer, can activate or inactivate the enzymatic nucleic acid domain of the nucleic acid sensor molecule, such that the activity of the nucleic acid sensor molecule is modulated in the presence of the target-signaling molecule. The nucleic acid sensor molecule can be designed to be active in the presence of the target molecule or alternately, can be designed to be inactive in the presence of the molecular target. For example, a nucleic acid sensor molecule is designed with a sensor domain having the sequence $(UUCA)_n$, where n is an integer from 1-10. In a non-limiting example, interaction of the HBV RT primer with the sensor domain of the nucleic acid sensor molecule can activate the enzymatic nucleic acid domain of the nucleic acid sensor molecule such that the sensor molecule catalyzes a reaction, for example cleavage of HBV RNA. In this example, the nucleic acid sensor molecule is activated in the presence of HBV RT or HBV RT primer, and can be used as a therapeutic to treat HBV infection. Alternately, the reaction can comprise cleavage or ligation of a labeled nucleic acid reporter molecule, providing a useful diagnostic reagent to detect the presence of HBV in a system.

HCV Target sites

Targets for useful nucleic acid molecules and nuclease activating compounds or chimeras can be determined as disclosed in Draper *et al.*, WO 93/23569; Sullivan *et al.*, WO 93/23057; Thompson *et al.*, WO 94/02595; Draper *et al.*, WO 95/04818; McSwiggen *et al.*, US Patent No. 5,525,468. Rather than repeat the guidance provided in those documents here, below are provided specific examples of such methods, not limiting to those in the art. Nucleic acid molecules and nuclease activating compounds or chimeras to such targets are designed as described in those applications and synthesized to be tested *in vitro* and *in vivo*, as also described. Such nucleic acid molecules and nuclease activating compounds or chimeras can also be optimized and delivered as described therein.

The sequence of HCV RNAs were screened for optimal enzymatic nucleic acid molecule target sites using a computer folding algorithm. Enzymatic nucleic acid cleavage sites were identified. These sites are shown in **Tables XVIII, XIX, XX and XXIII** (All sequences are 5' to 3' in the tables). The nucleotide base position is noted in the tables as that site to be cleaved by the designated type of enzymatic nucleic acid molecule. The nucleotide base position is noted in the tables as that site to be cleaved by the designated type of enzymatic nucleic acid molecule.

Because HCV RNAs are highly homologous in certain regions, some enzymatic nucleic acid molecule target sites are also homologous. In this case, a single enzymatic nucleic acid molecule will target different classes of HCV RNA. The advantage of one enzymatic nucleic acid molecule that targets several classes of HCV RNA is clear, especially in cases where one or more of these RNAs can contribute to the disease state.

Enzymatic nucleic acid molecules were designed that could bind and were individually analyzed by computer folding (Jaeger *et al.*, 1989 *Proc. Natl. Acad. Sci. USA*, 86, 7706) to assess whether the enzymatic nucleic acid molecule sequences fold into the appropriate secondary structure. Those enzymatic nucleic acid molecules with unfavorable intramolecular interactions between the binding arms and the catalytic core are eliminated from consideration. Varying binding arm lengths can be chosen to optimize activity. Generally, at least 5 bases on each arm are able to bind to, or otherwise interact with, the target RNA. Enzymatic nucleic acid molecules were designed to anneal to various sites in the mRNA message. The binding arms are complementary to the target site sequences described above.

HBV Target sites

Targets for useful ribozymes and antisense nucleic acids targeting HBV can be determined as disclosed in Draper *et al.*, WO 93/23569; Sullivan *et al.*, WO 93/23057; Thompson *et al.*, WO 94/02595; Draper *et al.*, WO 95/04818; McSwiggen *et al.*, US Patent No. 5,525,468. Other examples include the following PCT applications, which concern inactivation of expression of disease-related genes: WO 95/23225, WO 95/13380, WO 94/02595. Rather than repeat the guidance provided in those documents here, below are provided specific examples of such methods, not limiting to those in the art. Ribozymes and antisense to such targets are designed as described in those applications and synthesized to be tested *in vitro* and *in vivo*, as also described. The sequence of human HBV RNAs (for example, accession AF100308.1; HBV strain 2-18; additionally, other HBV strains can be screened by one skilled in the art, see **Table III** for other possible strains) were screened for optimal enzymatic nucleic acid and antisense target sites using a computer-folding algorithm. Antisense, hammerhead, DNAzyme, NCH (Inozyme), amberzyme, zinzyme or G-Cleaver ribozyme binding/cleavage sites were identified. These sites are shown in **Tables V to XI** (all sequences are 5' to 3' in the tables; X can be any base-paired sequence, the actual sequence is not relevant here). The nucleotide base position is noted in the Tables as that site to be cleaved by the designated type of enzymatic nucleic acid molecule. **Table IV** shows substrate positions selected from Renbo *et al.*, 1987, *Sci. Sin.*, 30, 507, used in Draper, USSN (07/882,712), filed May 14, 1992, entitled "METHOD AND REAGENT FOR INHIBITING HEPATITIS B VIRUS REPLICATION" and Draper et al., International PCT publication No. WO 93/23569, filed April 29, 1993, entitled "METHOD AND REAGENT FOR INHIBITING VIRAL REPLICATION". While human sequences can be screened and enzymatic nucleic acid molecule and/or antisense thereafter designed, as discussed in Stinchcomb *et al.*, WO 95/23225, mouse targeted ribozymes can be useful to test efficacy of action of the enzymatic nucleic acid molecule and/or antisense prior to testing in humans.

Antisense, hammerhead, DNAzyme, NCH (Inozyme), amberzyme, zinzyme or G-Cleaver ribozyme binding/cleavage sites were identified, as discussed above. The nucleic acid molecules were individually analyzed by computer folding (Jaeger *et al.*, 1989 *Proc. Natl. Acad. Sci. USA*, 86, 7706) to assess whether the sequences fold into the appropriate secondary structure. Those nucleic acid molecules with unfavorable intramolecular interactions such as between the binding arms and the catalytic core were eliminated from consideration. Varying binding arm lengths can be chosen to optimize activity.

Antisense, hammerhead, DNAzyme, NCH, amberzyme, zinzyme or G-Cleaver ribozyme binding/cleavage sites were identified and were designed to anneal to various sites in the RNA target. The binding arms are complementary to the target site sequences

described above. The nucleic acid molecules were chemically synthesized. The method of synthesis used follows the procedure for normal DNA/RNA synthesis as described below and in Usman *et al.*, 1987 *J. Am. Chem. Soc.*, 109, 7845; Scaringe *et al.*, 1990 *Nucleic Acids Res.*, 18, 5433; Wincott *et al.*, 1995 *Nucleic Acids Res.* 23, 2677-2684; and Caruthers *et al.*, 1992, *Methods in Enzymology* 211,3-19.

Synthesis of Nucleic acid Molecules

Synthesis of nucleic acids greater than 100 nucleotides in length is difficult using automated methods, and the therapeutic cost of such molecules is prohibitive. In this invention, small nucleic acid motifs ("small" refers to nucleic acid motifs no more than 100 nucleotides in length, preferably no more than 80 nucleotides in length, and most preferably no more than 50 nucleotides in length; *e.g.*, decoy nucleic acid molecules, aptamer nucleic acid molecules antisense nucleic acid molecules, enzymatic nucleic acid molecules) are preferably used for exogenous delivery. The simple structure of these molecules increases the ability of the nucleic acid to invade targeted regions of protein and/or RNA structure. Exemplary molecules of the instant invention are chemically synthesized, and others can similarly be synthesized.

Oligonucleotides (e.g., DNA oligonucleotides) are synthesized using protocols known in the art, for example as described in Caruthers *et al.*, 1992, *Methods in Enzymology* 211, 3-19, Thompson *et al.*, International PCT Publication No. WO 99/54459, Wincott *et al.*, 1995, *Nucleic Acids Res.* 23, 2677-2684, Wincott *et al.*, 1997, *Methods Mol. Bio.*, 74, 59, Brennan *et al.*, 1998, *Biotechnol Bioeng.*, 61, 33-45, and Brennan, US patent No. 6,001,311. The synthesis of oligonucleotides makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. In a non-limiting example, small scale syntheses are conducted on a 394 Applied Biosystems, Inc. synthesizer using a 0.2 μmol scale protocol with a 2.5 min coupling step for 2'-O-methylated nucleotides and a 45 sec coupling step for 2'-deoxy nucleotides. **Table II** outlines the amounts and the contact times of the reagents used in the synthesis cycle. Alternatively, syntheses at the 0.2 μmol scale can be performed on a 96-well plate synthesizer, such as the instrument produced by Protogene (Palo Alto, CA) with minimal modification to the cycle. A 33-fold excess (60 μL of 0.11 M = 6.6 μmol) of 2'-O-methyl phosphoramidite and a 105-fold excess of S-ethyl tetrazole (60 μL of 0.25 M = 15 μmol) can be used in each coupling cycle of 2'-O-methyl residues relative to polymer-bound 5'-hydroxyl. A 22-fold excess (40 μL of 0.11 M = 4.4 μmol) of deoxy phosphoramidite and a 70-fold excess of S-ethyl tetrazole (40 μL of 0.25 M = 10 μmol) can be used in each coupling cycle of deoxy residues relative to polymer-bound 5'-hydroxyl. Average coupling yields on the 394 Applied Biosystems, Inc. synthesizer, determined by colorimetric quantitation of the trityl fractions, are typically 97.5-

99%. Other oligonucleotide synthesis reagents for the 394 Applied Biosystems, Inc. synthesizer include the following: detritylation solution is 3% TCA in methylene chloride (ABI); capping is performed with 16% *N*-methyl imidazole in THF (ABI) and 10% acetic anhydride/10% 2,6-lutidine in THF (ABI); and oxidation solution is 16.9 mM I_2, 49 mM pyridine, 9% water in THF (PERSEPTIVE™). Burdick & Jackson Synthesis Grade acetonitrile is used directly from the reagent bottle. S-Ethyltetrazole solution (0.25 M in acetonitrile) is made up from the solid obtained from American International Chemical, Inc. Alternately, for the introduction of phosphorothioate linkages, Beaucage reagent (3H-1,2-Benzodithiol-3-one 1,1-dioxide, 0.05 M in acetonitrile) is used.

Deprotection of the DNA-based oligonucleotides is performed as follows: the polymer-bound trityl-on oligoribonucleotide is transferred to a 4 mL glass screw top vial and suspended in a solution of 40% aq. methylamine (1 mL) at 65 °C for 10 min. After cooling to –20 °C, the supernatant is removed from the polymer support. The support is washed three times with 1.0 mL of EtOH:MeCN:H2O/3:1:1, vortexed and the supernatant is then added to the first supernatant. The combined supernatants, containing the oligoribonucleotide, are dried to a white powder.

The method of synthesis used for normal RNA including certain decoy nucleic acid molecules and enzymatic nucleic acid molecules follows the procedure as described in Usman *et al.*, 1987, *J. Am. Chem. Soc.*, 109, 7845; Scaringe *et al.*, 1990, *Nucleic Acids Res.*, 18, 5433; and Wincott *et al.*, 1995, *Nucleic Acids Res.* 23, 2677-2684 Wincott *et al.*, 1997, *Methods Mol. Bio.*, 74, 59, and makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. In a non-limiting example, small scale syntheses are conducted on a 394 Applied Biosystems, Inc. synthesizer using a 0.2 µmol scale protocol with a 7.5 min coupling step for alkylsilyl protected nucleotides and a 2.5 min coupling step for 2'-O-methylated nucleotides. **Table II** outlines the amounts and the contact times of the reagents used in the synthesis cycle. Alternatively, syntheses at the 0.2 µmol scale can be done on a 96-well plate synthesizer, such as the instrument produced by Protogene (Palo Alto, CA) with minimal modification to the cycle. A 33-fold excess (60 µL of 0.11 M = 6.6 µmol) of 2'-O-methyl phosphoramidite and a 75-fold excess of S-ethyl tetrazole (60 µL of 0.25 M = 15 µmol) can be used in each coupling cycle of 2'-O-methyl residues relative to polymer-bound 5'-hydroxyl. A 66-fold excess (120 µL of 0.11 M = 13.2 µmol) of alkylsilyl (ribo) protected phosphoramidite and a 150-fold excess of S-ethyl tetrazole (120 µL of 0.25 M = 30 µmol) can be used in each coupling cycle of ribo residues relative to polymer-bound 5'-hydroxyl. Average coupling yields on the 394 Applied Biosystems, Inc. synthesizer, determined by colorimetric quantitation of the trityl fractions, are typically 97.5-99%. Other oligonucleotide synthesis reagents for the 394 Applied Biosystems, Inc. synthesizer include the following: detritylation

solution is 3% TCA in methylene chloride (ABI); capping is performed with 16% *N*-methyl imidazole in THF (ABI) and 10% acetic anhydride/10% 2,6-lutidine in THF (ABI); oxidation solution is 16.9 mM I_2, 49 mM pyridine, 9% water in THF (PERSEPTIVE™). Burdick & Jackson Synthesis Grade acetonitrile is used directly from the reagent bottle. S-Ethyltetrazole solution (0.25 M in acetonitrile) is made up from the solid obtained from American International Chemical, Inc. Alternately, for the introduction of phosphorothioate linkages, Beaucage reagent (3H-1,2-Benzodithiol-3-one 1,1-dioxide0.05 M in acetonitrile) is used.

Deprotection of the RNA is performed using either a two-pot or one-pot protocol. For the two-pot protocol, the polymer-bound trityl-on oligoribonucleotide is transferred to a 4 mL glass screw top vial and suspended in a solution of 40% aq. methylamine (1 mL) at 65 °C for 10 min. After cooling to –20 °C, the supernatant is removed from the polymer support. The support is washed three times with 1.0 mL of EtOH:MeCN:H2O/3:1:1, vortexed and the supernatant is then added to the first supernatant. The combined supernatants, containing the oligoribonucleotide, are dried to a white powder. The base deprotected oligoribonucleotide is resuspended in anhydrous TEA/HF/NMP solution (300 µL of a solution of 1.5 mL N-methylpyrrolidinone, 750 µL TEA and 1 mL TEA•3HF to provide a 1.4 M HF concentration) and heated to 65 °C. After 1.5 h, the oligomer is quenched with 1.5 M NH_4HCO_3.

Alternatively, for the one-pot protocol, the polymer-bound trityl-on oligoribonucleotide is transferred to a 4 mL glass screw top vial and suspended in a solution of 33% ethanolic methylamine/DMSO: 1/1 (0.8 mL) at 65 °C for 15 min. The vial is brought to r.t. TEA•3HF (0.1 mL) is added and the vial is heated at 65 °C for 15 min. The sample is cooled at –20 °C and then quenched with 1.5 M NH_4HCO_3.

For purification of the trityl-on oligomers, the quenched NH_4HCO_3 solution is loaded onto a C-18 containing cartridge that had been prewashed with acetonitrile followed by 50 mM TEAA. After washing the loaded cartridge with water, the RNA is detritylated with 0.5% TFA for 13 min. The cartridge is then washed again with water, salt exchanged with 1 M NaCl and washed with water again. The oligonucleotide is then eluted with 30% acetonitrile.

Inactive hammerhead ribozymes or binding attenuated control (BAC) oligonucleotides are synthesized by substituting a U for G_5 and a U for A_{14} (numbering from Hertel, K. J., *et al.*, 1992, *Nucleic Acids Res.*, 20, 3252). Similarly, one or more nucleotide substitutions can be introduced in other nucleic acid decoy molecules to inactivate the molecule and such molecules can serve as a negative control.

The average stepwise coupling yields are typically >98% (Wincott *et al.*, 1995 *Nucleic Acids Res.* 23, 2677-2684). Those of ordinary skill in the art will recognize that the scale of synthesis can be adapted to be larger or smaller than the example described above including but not limited to 96-well format, all that is important is the ratio of chemicals used in the reaction.

Alternatively, the nucleic acid molecules of the present invention can be synthesized separately and joined together post-synthetically, for example, by ligation (Moore *et al.*, 1992, *Science* 256, 9923; Draper *et al.*, International PCT publication No. WO 93/23569; Shabarova *et al.*, 1991, *Nucleic Acids Research* 19, 4247; Bellon *et al.*, 1997, *Nucleosides & Nucleotides*, 16, 951; Bellon *et al.*, 1997, *Bioconjugate Chem.* 8, 204).

The nucleic acid molecules of the present invention can be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-*C*-allyl, 2'-flouro, 2'-*O*-methyl, 2'-H (for a review see Usman and Cedergren, 1992, *TIBS* 17, 34; Usman *et al.*, 1994, *Nucleic Acids Symp. Ser.* 31, 163). Ribozymes can be purified by gel electrophoresis using general methods or can be purified by high pressure liquid chromatography (HPLC; see Wincott *et al.*, *supra*, the totality of which is hereby incorporated herein by reference) and re-suspended in water.

The sequences of the nucleic acid molecules that are chemically synthesized, useful in this study, are shown in **Tables XI, XV, XX, XXI, XXII and XXIII**. The nucleic acid sequences listed in **Tables IV-XI, XIV-XV and XVIII-XXIII** can be formed of ribonucleotides or other nucleotides or non-nucleotides. Such nucleic acid sequences are equivalent to the sequences described specifically in the Tables.

Optimizing Activity of the nucleic acid molecule of the invention

Chemically synthesizing nucleic acid molecules with modifications (base, sugar and/or phosphate) can prevent their degradation by serum ribonucleases, which can increase their potency (see *e.g.*, Eckstein *et al.*, International Publication No. WO 92/07065; Perrault *et al.*, 1990 *Nature* 344, 565; Pieken et al., 1991, *Science* 253, 314; Usman and Cedergren, 1992, *Trends in Biochem. Sci.* 17, 334; Usman *et al.*, International Publication No. WO 93/15187; and Rossi *et al.*, International Publication No. WO 91/03162; Sproat, US Patent No. 5,334,711; Gold *et al.*, US 6,300,074; and Burgin *et al., supra;* all of which are incorporated by reference herein). All of the above references describe various chemical modifications that can be made to the base, phosphate and/or sugar moieties of the nucleic acid molecules described herein. Modifications that enhance their efficacy in cells, and removal of bases from nucleic acid molecules to shorten oligonucleotide synthesis times and reduce chemical requirements are desired.

There are several examples in the art describing sugar, base and phosphate modifications that can be introduced into nucleic acid molecules with significant enhancement in their nuclease stability and efficacy. For example, oligonucleotides are modified to enhance stability and/or enhance biological activity by modification with nuclease resistant groups, for example, 2'-amino, 2'-*C*-allyl, 2'-flouro, 2'-*O*-methyl, 2'-H, nucleotide base modifications (for a review see Usman and Cedergren, 1992, *TIBS*. 17, 34; Usman *et al.,* 1994, *Nucleic Acids Symp. Ser.* 31, 163; Burgin *et al.*, 1996, *Biochemistry*, 35, 14090). Sugar modification of nucleic acid molecules have been extensively described in the art (see Eckstein *et al.*, *International Publication* PCT No. WO 92/07065; Perrault *et al. Nature*, 1990, *344*, 565-568; Pieken *et al. Science,* 1991, *253*, 314-317; Usman and Cedergren, *Trends in Biochem. Sci.* , 1992, *17*, 334-339; Usman *et al. International Publication* PCT No. WO 93/15187; Sproat, *US Patent* No. 5,334,711 and Beigelman *et al.*, 1995, *J. Biol. Chem.*, 270, 25702; Beigelman *et al.*, International PCT publication No. WO 97/26270; Beigelman *et al.*, US Patent No. 5,716,824; Usman *et al.*, US patent No. 5,627,053; Woolf *et al.*, International PCT Publication No. WO 98/13526; Thompson *et al.*, USSN 60/082,404 which was filed on April 20, 1998; Karpeisky *et al.*, 1998, *Tetrahedron Lett.*, 39, 1131; Earnshaw and Gait, 1998, *Biopolymers (Nucleic Acid Sciences)*, 48, 39-55; Verma and Eckstein, 1998, *Annu. Rev. Biochem.*, 67, 99-134; and Burlina *et al.*, 1997, *Bioorg. Med. Chem.*, 5, 1999-2010; all of the references are hereby incorporated in their totality by reference herein). Such publications describe general methods and strategies to determine the location of incorporation of sugar, base and/or phosphate modifications and the like into ribozymes without modulating catalysis, and are incorporated by reference herein. In view of such teachings, similar modifications can be used as described herein to modify the nucleic acid molecules of the instant invention.

While chemical modification of oligonucleotide internucleotide linkages with phosphorothioate, phosphorothioate, and/or 5'-methylphosphonate linkages improves stability, excessive modifications can cause some toxicity. Therefore, when designing nucleic acid molecules, the amount of these internucleotide linkages should be minimized. The reduction in the concentration of these linkages should lower toxicity, resulting in increased efficacy and higher specificity of these molecules.

Nucleic acid molecules having chemical modifications that maintain or enhance activity are provided. Such a nucleic acid is also generally more resistant to nucleases than an unmodified nucleic acid. Accordingly, the *in vitro* and/or *in vivo* activity should not be significantly lowered. In cases in which modulation is the goal, therapeutic nucleic acid molecules delivered exogenously should optimally be stable within cells until translation of the target RNA has been modulated long enough to reduce the levels of the undesirable protein. This period of time varies between hours to days depending upon the disease state.

Improvements in the chemical synthesis of RNA and DNA (Wincott *et al.*, 1995 *Nucleic Acids Res.* 23, 2677; Caruthers *et al.*, 1992, *Methods in Enzymology* 211,3-19 (incorporated by reference herein)) have expanded the ability to modify nucleic acid molecules by introducing nucleotide modifications to enhance their nuclease stability, as described above.

In one embodiment, nucleic acid molecules of the invention include one or more G-clamp nucleotides. A G-clamp nucleotide is a modified cytosine analog wherein the modifications confer the ability to hydrogen bond both Watson-Crick and Hoogsteen faces of a complementary guanine within a duplex, see for example Lin and Matteucci, 1998, *J. Am. Chem. Soc.,* 120, 8531-8532. A single G-clamp analog substation within an oligonucleotide can result in substantially enhanced helical thermal stability and mismatch discrimination when hybridized to complementary oligonucleotides. The inclusion of such nucleotides in nucleic acid molecules of the invention results in both enhanced affinity and specificity to nucleic acid targets. In another embodiment, nucleic acid molecules of the invention include one or more LNA "locked nucleic acid" nucleotides such as a 2', 4'-C methylene bicyclo nucleotide (see for example Wengel *et al.,* International PCT Publication No. WO 00/66604 and WO 99/14226).

In another embodiment, the invention features conjugates and/or complexes of nucleic acid molecules targeting HBV or HCV. Such conjugates and/or complexes can be used to facilitate delivery of molecules into a biological system, such as a cell. The conjugates and complexes provided by the instant invention can impart therapeutic activity by transferring therapeutic compounds across cellular membranes, altering the pharmacokinetics, and/or modulating the localization of nucleic acid molecules of the invention. The present invention encompasses the design and synthesis of novel conjugates and complexes for the delivery of molecules, including, but not limited to, small molecules, lipids, phospholipids, nucleosides, nucleotides, nucleic acids, antibodies, toxins, negatively charged polymers and other polymers, for example proteins, peptides, hormones, carbohydrates, polyethylene glycols, or polyamines, across cellular membranes. In general, the transporters described are designed to be used either individually or as part of a multi-component system, with or without degradable linkers. These compounds are expected to improve delivery and/or localization of nucleic acid molecules of the invention into a number of cell types originating from different tissues, in the presence or absence of serum (see Sullenger and Cech, US 5,854,038). Conjugates of the molecules described herein can be attached to biologically active molecules via linkers that are biodegradable, such as biodegradable nucleic acid linker molecules.

The term "biodegradable nucleic acid linker molecule" as used herein, refers to a nucleic acid molecule that is designed as a biodegradable linker to connect one molecule to another molecule, for example, a biologically active molecule. The stability of the

biodegradable nucleic acid linker molecule can be modulated by using various combinations of ribonucleotides, deoxyribonucleotides, and chemically modified nucleotides, for example, 2'-O-methyl, 2'-fluoro, 2'-amino, 2'-O-amino, 2'-C-allyl, 2'-O-allyl, and other 2'-modified or base modified nucleotides. The biodegradable nucleic acid linker molecule can be a dimer, trimer, tetramer or longer nucleic acid molecule, for example, an oligonucleotide of about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides in length, or can comprise a single nucleotide with a phosphorus-based linkage, for example, a phosphoramidate or phosphodiester linkage. The biodegradable nucleic acid linker molecule can also comprise nucleic acid backbone, nucleic acid sugar, or nucleic acid base modifications.

The term "biodegradable" as used herein, refers to degradation in a biological system, for example enzymatic degradation or chemical degradation.

The term "biologically active molecule" as used herein, refers to compounds or molecules that are capable of eliciting or modifying a biological response in a system. Non-limiting examples of biologically active molecules contemplated by the instant invention include therapeutically active molecules such as antibodies, hormones, antivirals, peptides, proteins, chemotherapeutics, small molecules, vitamins, co-factors, nucleosides, nucleotides, oligonucleotides, enzymatic nucleic acids, antisense nucleic acids, triplex forming oligonucleotides, 2,5-A chimeras, siRNA, dsRNA, allozymes, aptamers, decoys and analogs thereof. Biologically active molecules of the invention also include molecules capable of modulating the pharmacokinetics and/or pharmacodynamics of other biologically active molecules, for example, lipids and polymers such as polyamines, polyamides, polyethylene glycol and other polyethers.

The term "phospholipid" as used herein, refers to a hydrophobic molecule comprising at least one phosphorus group. For example, a phospholipid can comprise a phosphorus-containing group and saturated or unsaturated alkyl group, optionally substituted with OH, COOH, oxo, amine, or substituted or unsubstituted aryl groups.

Therapeutic nucleic acid molecules (*e.g.,* decoy nucleic acid molecules) delivered exogenously optimally are stable within cells until reverse trascription of the pregenomic RNA has been modulated long enough to reduce the levels of HBV or HCV DNA. The nucleic acid molecules are resistant to nucleases in order to function as effective intracellular therapeutic agents. Improvements in the chemical synthesis of nucleic acid molecules described in the instant invention and in the art have expanded the ability to modify nucleic acid molecules by introducing nucleotide modifications to enhance their nuclease stability as described above.

In yet another embodiment, nucleic acid molecules having chemical modifications that maintain or enhance enzymatic activity are provided. Such nucleic acids are also generally more resistant to nucleases than unmodified nucleic acids. Thus, *in vitro* and/or *in vivo* the activity should not be significantly lowered. As exemplified herein, such nucleic acid molecules are useful *in vitro* and/or *in vivo* even if activity over all is reduced 10 fold (Burgin *et al.*, 1996, *Biochemistry*, 35, 14090).

Use of the nucleic acid-based molecules of the invention will lead to better treatment of the disease progression by affording the possibility of combination therapies (*e.g.*, multiple antisense, nucleic acid decoy, or nucleic acid aptamer molecules targeted to different genes; nucleic acid molecules coupled with known small molecule modulators ors; or intermittent treatment with combinations of molecules (including different motifs) and/or other chemical or biological molecules). The treatment of patients with nucleic acid molecules may also include combinations of different types of nucleic acid molecules.

In another aspect the nucleic acid molecules comprise a 5' and/or a 3'- cap structure.

By "cap structure" is meant chemical modifications, which have been incorporated at either terminus of the oligonucleotide (see, for example, Wincott *et al.*, WO 97/26270, incorporated by reference herein). These terminal modifications protect the nucleic acid molecule from exonuclease degradation, and may help in delivery and/or localization within a cell. The cap may be present at the 5'-terminus (5'-cap) or at the 3'-terminal (3'-cap) or may be present on both termini. In non-limiting examples: the 5'-cap is selected from the group comprising inverted abasic residue (moiety); 4',5'-methylene nucleotide; 1-(beta-D-erythrofuranosyl) nucleotide, 4'-thio nucleotide; carbocyclic nucleotide; 1,5-anhydrohexitol nucleotide; L-nucleotides; alpha-nucleotides; modified base nucleotide; phosphorodithioate linkage; *threo*-pentofuranosyl nucleotide; acyclic 3',4'-seco nucleotide; acyclic 3,4-dihydroxybutyl nucleotide; acyclic 3,5-dihydroxypentyl nucleotide, 3'-3'-inverted nucleotide moiety; 3'-3'-inverted abasic moiety; 3'-2'-inverted nucleotide moiety; 3'-2'-inverted abasic moiety; 1,4-butanediol phosphate; 3'-phosphoramidate; hexylphosphate; aminohexyl phosphate; 3'-phosphate; 3'-phosphorothioate; phosphorodithioate; or bridging or non-bridging methylphosphonate moiety (for more details, see Wincott *et al.*, International PCT publication No. WO 97/26270, incorporated by reference herein).

In yet another preferred embodiment, the 3'-cap is selected from a group comprising, 4',5'-methylene nucleotide; 1-(beta-D-erythrofuranosyl) nucleotide; 4'-thio nucleotide, carbocyclic nucleotide; 5'-amino-alkyl phosphate; 1,3-diamino-2-propyl phosphate; 3-aminopropyl phosphate; 6-aminohexyl phosphate; 1,2-aminododecyl phosphate; hydroxypropyl phosphate; 1,5-anhydrohexitol nucleotide; L-nucleotide; alpha-nucleotide; modified base nucleotide; phosphorodithioate; *threo*-pentofuranosyl nucleotide; acyclic 3',4'-

seco nucleotide; 3,4-dihydroxybutyl nucleotide; 3,5-dihydroxypentyl nucleotide, 5'-5'-inverted nucleotide moiety; 5'-5'-inverted abasic moiety; 5'-phosphoramidate; 5'-phosphorothioate; 1,4-butanediol phosphate; 5'-amino; bridging and/or non-bridging 5'-phosphoramidate, phosphorothioate and/or phosphorodithioate, bridging or non bridging methylphosphonate and 5'-mercapto moieties (for more details see Beaucage and Iyer, 1993, *Tetrahedron* 49, 1925; incorporated by reference herein).

By the term "non-nucleotide" is meant any group or compound which can be incorporated into a nucleic acid chain in the place of one or more nucleotide units, including either sugar and/or phosphate substitutions, and allows the remaining bases to exhibit their enzymatic activity. The group or compound is abasic in that it does not contain a commonly recognized nucleotide base, such as adenosine, guanine, cytosine, uracil or thymine.

The term "alkyl" as used herein refers to a saturated aliphatic hydrocarbon, including straight-chain, branched-chain "isoalkyl", and cyclic alkyl groups. The term "alkyl" also comprises alkoxy, alkyl-thio, alkyl-thio-alkyl, alkoxyalkyl, alkylamino, alkenyl, alkynyl, alkoxy, cycloalkenyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heteroaryl, C1-C6 hydrocarbyl, aryl or substituted aryl groups. Preferably, the alkyl group has 1 to 12 carbons. More preferably it is a lower alkyl of from about 1 to 7 carbons, more preferably about 1 to 4 carbons. The alkyl group can be substituted or unsubstituted. When substituted the substituted group(s) preferably comprise hydroxy, oxy, thio, amino, nitro, cyano, alkoxy, alkyl-thio, alkyl-thio-alkyl, alkoxyalkyl, alkylamino, silyl, alkenyl, alkynyl, alkoxy, cycloalkenyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heteroaryl, C1-C6 hydrocarbyl, aryl or substituted aryl groups. The term "alkyl" also includes alkenyl groups containing at least one carbon-carbon double bond, including straight-chain, branched-chain, and cyclic groups. Preferably, the alkenyl group has about 2 to 12 carbons. More preferably it is a lower alkenyl of from about 2 to 7 carbons, more preferably about 2 to 4 carbons. The alkenyl group can be substituted or unsubstituted. When substituted the substituted group(s) preferably comprise hydroxy, oxy, thio, amino, nitro, cyano, alkoxy, alkyl-thio, alkyl-thio-alkyl, alkoxyalkyl, alkylamino, silyl, alkenyl, alkynyl, alkoxy, cycloalkenyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heteroaryl, C1-C6 hydrocarbyl, aryl or substituted aryl groups. The term "alkyl" also includes alkynyl groups containing at least one carbon-carbon triple bond, including straight-chain, branched-chain, and cyclic groups. Preferably, the alkynyl group has about 2 to 12 carbons. More preferably it is a lower alkynyl of from about 2 to 7 carbons, more preferably about 2 to 4 carbons. The alkynyl group can be substituted or unsubstituted. When substituted the substituted group(s) preferably comprise hydroxy, oxy, thio, amino, nitro, cyano, alkoxy, alkyl-thio, alkyl-thio-alkyl, alkoxyalkyl, alkylamino, silyl, alkenyl, alkynyl, alkoxy, cycloalkenyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heteroaryl, C1-C6 hydrocarbyl, aryl or substituted aryl groups. Alkyl groups or moieties of

the invention can also include aryl, alkylaryl, carbocyclic aryl, heterocyclic aryl, amide and ester groups. The preferred substituent(s) of aryl groups are halogen, trihalomethyl, hydroxyl, SH, OH, cyano, alkoxy, alkyl, alkenyl, alkynyl, and amino groups. An "alkylaryl" group refers to an alkyl group (as described above) covalently joined to an aryl group (as described above). Carbocyclic aryl groups are groups wherein the ring atoms on the aromatic ring are all carbon atoms. The carbon atoms are optionally substituted. Heterocyclic aryl groups are groups having from about 1 to 3 heteroatoms as ring atoms in the aromatic ring and the remainder of the ring atoms are carbon atoms. Suitable heteroatoms include oxygen, sulfur, and nitrogen, and include furanyl, thienyl, pyridyl, pyrrolyl, N-lower alkyl pyrrolo, pyrimidyl, pyrazinyl, imidazolyl and the like, all optionally substituted. An "amide" refers to an -C(O)-NH-R, where R is either alkyl, aryl, alkylaryl or hydrogen. An "ester" refers to an -C(O)-OR', where R is either alkyl, aryl, alkylaryl or hydrogen.

The term "alkoxyalkyl" as used herein refers to an alkyl-O-alkyl ether, for example methoxyethyl or ethoxymethyl.

The term "alkyl-thio-alkyl" as used herein refers to an alkyl-S-alkyl thioether, for example methylthiomethyl or methylthioethyl.

The term "amination" as used herein refers to a process in which an amino group or substituted amine is introduced into an organic molecule.

The term "exocyclic amine protecting moiety" as used herein refers to a nucleobase amino protecting group compatible with oligonucleotide synthesis, for example an acyl or amide group.

The term "alkenyl" as used herein refers to a straight or branched hydrocarbon of a designed number of carbon atoms containing at least one carbon-carbon double bond. Examples of "alkenyl" include vinyl, allyl, and 2-methyl-3-heptene.

The term "alkoxy" as used herein refers to an alkyl group of indicated number of carbon atoms attached to the parent molecular moiety through an oxygen bridge. Examples of alkoxy groups include, for example, methoxy, ethoxy, propoxy and isopropoxy.

The term "alkynyl" as used herein refers to a straight or branched hydrocarbon of a designed number of carbon atoms containing at least one carbon-carbon triple bond. Examples of "alkynyl" include propargyl, propyne, and 3-hexyne.

The term "aryl" as used herein refers to an aromatic hydrocarbon ring system containing at least one aromatic ring. The aromatic ring can optionally be fused or otherwise attached to other aromatic hydrocarbon rings or non-aromatic hydrocarbon rings. Examples

of aryl groups include, for example, phenyl, naphthyl, 1,2,3,4-tetrahydronaphthalene and biphenyl. Preferred examples of aryl groups include phenyl and naphthyl.

The term "cycloalkenyl" as used herein refers to a C3-C8 cyclic hydrocarbon containing at least one carbon-carbon double bond. Examples of cycloalkenyl include cyclopropenyl, cyclobutenyl, cyclopentenyl, cyclopentadiene, cyclohexenyl, 1,3-cyclohexadiene, cycloheptenyl, cycloheptatrienyl, and cyclooctenyl.

The term "cycloalkyl" as used herein refers to a C3-C8 cyclic hydrocarbon. Examples of cycloalkyl include cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl and cyclooctyl.

The term "cycloalkylalkyl," as used herein, refers to a C3-C7 cycloalkyl group attached to the parent molecular moiety through an alkyl group, as defined above. Examples of cycloalkylalkyl groups include cyclopropylmethyl and cyclopentylethyl.

The terms "halogen" or "halo" as used herein refers to indicate fluorine, chlorine, bromine, and iodine.

The term "heterocycloalkyl," as used herein refers to a non-aromatic ring system containing at least one heteroatom selected from nitrogen, oxygen, and sulfur. The heterocycloalkyl ring can be optionally fused to or otherwise attached to other heterocycloalkyl rings and/or non-aromatic hydrocarbon rings. Preferred heterocycloalkyl groups have from 3 to 7 members. Examples of heterocycloalkyl groups include, for example, piperazine, morpholine, piperidine, tetrahydrofuran, pyrrolidine, and pyrazole. Preferred heterocycloalkyl groups include piperidinyl, piperazinyl, morpholinyl, and pyrolidinyl.

The term "heteroaryl" as used herein refers to an aromatic ring system containing at least one heteroatom selected from nitrogen, oxygen, and sulfur. The heteroaryl ring can be fused or otherwise attached to one or more heteroaryl rings, aromatic or non-aromatic hydrocarbon rings or heterocycloalkyl rings. Examples of heteroaryl groups include, for example, pyridine, furan, thiophene, 5,6,7,8-tetrahydroisoquinoline and pyrimidine. Preferred examples of heteroaryl groups include thienyl, benzothienyl, pyridyl, quinolyl, pyrazinyl, pyrimidyl, imidazolyl, benzimidazolyl, furanyl, benzofuranyl, thiazolyl, benzothiazolyl, isoxazolyl, oxadiazolyl, isothiazolyl, benzisothiazolyl, triazolyl, tetrazolyl, pyrrolyl, indolyl, pyrazolyl, and benzopyrazolyl.

The term "C1-C6 hydrocarbyl" as used herein refers to straight, branched, or cyclic alkyl groups having 1-6 carbon atoms, optionally containing one or more carbon-carbon double or triple bonds. Examples of hydrocarbyl groups include, for example, methyl, ethyl,

propyl, isopropyl, n-butyl, sec-butyl, tert-butyl, pentyl, 2-pentyl, isopentyl, neopentyl, hexyl, 2-hexyl, 3-hexyl, 3-methylpentyl, vinyl, 2-pentene, cyclopropylmethyl, cyclopropyl, cyclohexylmethyl, cyclohexyl and propargyl. When reference is made herein to C1-C6 hydrocarbyl containing one or two double or triple bonds it is understood that at least two carbons are present in the alkyl for one double or triple bond, and at least four carbons for two double or triple bonds.

The term "nucleotide" as used herein refers to a heterocyclic nitrogenous base in N-glycosidic linkage with a phosphorylated sugar. Nucleotides are recognized in the art to include natural bases (standard), and modified bases well known in the art. Such bases are generally located at the 1' position of a nucleotide sugar moiety. Nucleotides generally comprise a base, sugar and a phosphate group. The nucleotides can be unmodified or modified at the sugar, phosphate and/or base moiety, (also referred to interchangeably as nucleotide analogs, modified nucleotides, non-natural nucleotides, non-standard nucleotides and other; see for example, Usman and McSwiggen, *supra*; Eckstein *et al.*, International PCT Publication No. WO 92/07065; Usman *et al.*, International PCT Publication No. WO 93/15187; Uhlman & Peyman, supra all are hereby incorporated by reference herein. There are several examples of modified nucleic acid bases known in the art as summarized by Limbach *et al.*, 1994, Nucleic Acids Res. 22, 2183. Some of the non-limiting examples of chemically modified and other natural nucleic acid bases that can be introduced into nucleic acids include, for example, inosine, purine, pyridin-4-one, pyridin-2-one, phenyl, pseudouracil, 2, 4, 6-trimethoxy benzene, 3-methyl uracil, dihydrouridine, naphthyl, aminophenyl, 5-alkylcytidines (*e.g.*, 5-methylcytidine), 5-alkyluridines (*e.g.*, ribothymidine), 5-halouridine (*e.g.*, 5-bromouridine) or 6-azapyrimidines or 6-alkylpyrimidines (*e.g.* 6-methyluridine), propyne, quesosine, 2-thiouridine, 4-thiouridine, wybutosine, wybutoxosine, 4-acetylcytidine, 5-(carboxyhydroxymethyl)uridine, 5'-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluridine, beta-D-galactosylqueosine, 1-methyladenosine, 1-methylinosine, 2,2-dimethylguanosine, 3-methylcytidine, 2-methyladenosine, 2-methylguanosine, N6-methyladenosine, 7-methylguanosine, 5-methoxyaminomethyl-2-thiouridine, 5-methylaminomethyluridine, 5-methylcarbonylmethyluridine, 5-methyloxyuridine, 5-methyl-2-thiouridine, 2-methylthio-N6-isopentenyladenosine, beta-D-mannosylqueosine, uridine-5-oxyacetic acid, 2-thiocytidine, threonine derivatives and others (Burgin *et al.*, 1996, Biochemistry, 35, 14090; Uhlman & Peyman, *supra*). By "modified bases" in this aspect is meant nucleotide bases other than adenine, guanine, cytosine and uracil at 1' position or their equivalents; such bases can be used at any position, for example, within the catalytic core of an enzymatic nucleic acid molecule and/or in the substrate-binding regions of the nucleic acid molecule.

The term "nucleoside" as used herein refers to a heterocyclic nitrogenous base in N-glycosidic linkage with a sugar. Nucleosides are recognized in the art to include natural bases (standard), and modified bases well known in the art. Such bases are generally located at the 1' position of a nucleoside sugar moiety. Nucleosides generally comprise a base and sugar group. The nucleosides can be unmodified or modified at the sugar, and/or base moiety (also referred to interchangeably as nucleoside analogs, modified nucleosides, non-natural nucleosides, non-standard nucleosides and other; see for example, Usman and McSwiggen, *supra*; Eckstein *et al.*, International PCT Publication No. WO 92/07065; Usman *et al.*, International PCT Publication No. WO 93/15187; Uhlman & Peyman, supra all are hereby incorporated by reference herein). There are several examples of modified nucleic acid bases known in the art as summarized by Limbach *et al.*, 1994, Nucleic Acids Res. 22, 2183. Some of the non-limiting examples of chemically modified and other natural nucleic acid bases that can be introduced into nucleic acids include, inosine, purine, pyridin-4-one, pyridin-2-one, phenyl, pseudouracil, 2, 4, 6-trimethoxy benzene, 3-methyl uracil, dihydrouridine, naphthyl, aminophenyl, 5-alkylcytidines (*e.g.*, 5-methylcytidine), 5-alkyluridines (*e.g.*, ribothymidine), 5-halouridine (*e.g.*, 5-bromouridine) or 6-azapyrimidines or 6-alkylpyrimidines (*e.g.* 6-methyluridine), propyne, quesosine, 2-thiouridine, 4-thiouridine, wybutosine, wybutoxosine, 4-acetylcytidine, 5-(carboxyhydroxymethyl)uridine, 5'-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluridine, beta-D-galactosylqueosine, 1-methyladenosine, 1-methylinosine, 2,2-dimethylguanosine, 3-methylcytidine, 2-methyladenosine, 2-methylguanosine, N6-methyladenosine, 7-methylguanosine, 5-methoxyaminomethyl-2-thiouridine, 5-methylaminomethyluridine, 5-methylcarbonylmethyluridine, 5-methyloxyuridine, 5-methyl-2-thiouridine, 2-methylthio-N6-isopentenyladenosine, beta-D-mannosylqueosine, uridine-5-oxyacetic acid, 2-thiocytidine, threonine derivatives and others (Burgin *et al.*, 1996, Biochemistry, 35, 14090; Uhlman & Peyman, *supra*). By "modified bases" in this aspect is meant nucleoside bases other than adenine, guanine, cytosine and uracil at 1' position or their equivalents; such bases can be used at any position, for example, within the catalytic core of an enzymatic nucleic acid molecule and/or in the substrate-binding regions of the nucleic acid molecule.

In one embodiment, the invention features modified nucleic acid molecules with phosphate backbone modifications comprising one or more phosphorothioate, phosphorodithioate, methylphosphonate, morpholino, amidate carbamate, carboxymethyl, acetamidate, polyamide, sulfonate, sulfonamide, sulfamate, formacetal, thioformacetal, and/or alkylsilyl, substitutions. For a review of oligonucleotide backbone modifications see Hunziker and Leumann, 1995, *Nucleic Acid Analogues: Synthesis and Properties,* in *Modern Synthetic Methods,* VCH, 331-417, and Mesmaeker *et al.,* 1994, *Novel Backbone Replacements for Oligonucleotides,* in *Carbohydrate Modifications in Antisense Research,* ACS, 24-39. These references are hereby incorporated by reference herein.

The term "abasic" as used herein refers to sugar moieties lacking a base or having other chemical groups in place of a base at the 1' position, for example a 3',3'-linked or 5',5'-linked deoxyabasic ribose derivative (for more details see Wincott *et al.,* International PCT publication No. WO 97/26270).

The term "unmodified nucleoside" as used herein refers to one of the bases adenine, cytosine, guanine, thymine, uracil joined to the 1' carbon of β-D-ribo-furanose.

The term "modified nucleoside" as used herein refers to any nucleotide base which contains a modification in the chemical structure of an unmodified nucleotide base, sugar and/or phosphate.

In connection with 2'-modified nucleotides as described for the present invention, by "amino" is meant 2'-NH_2 or 2'-*O*- NH_2, which can be modified or unmodified. Such modified groups are described, for example, in Eckstein *et al.*, U.S. Patent 5,672,695 and Matulic-Adamic *et al.,* WO 98/28317, respectively, which are both incorporated by reference in their entireties.

Various modifications to nucleic acid (*e.g.*, enzymatic nucleic acid, antisense, decoy, aptamer, siRNA, triplex oligonucleotides, 2,5-A oligonucleotides and other nucleic acid molecules) structure can be made to enhance the utility of these molecules. For example, such modifications can enhance shelf life, half-life *in vitro,* stability, and ease of introduction of such oligonucleotides to the target site, including *e.g.*, enhancing penetration of cellular membranes and conferring the ability to recognize and bind to targeted cells.

Use of these molecules can lead to better treatment of the disease progression by affording the possibility of combination therapies (*e.g.*, multiple nucleic acid molecules targeted to different genes, nucleic acid molecules coupled with known small molecule inhibitors, or intermittent treatment with combinations of nucleic acid molecules (including different nucleic acid molecule motifs) and/or other chemical or biological molecules). The treatment of patients with nucleic acid molecules can also include combinations of different types of nucleic acid molecules. Therapies can be devised which include a mixture of enzymatic nucleic acid molecules (including different enzymatic nucleic acid molecule motifs), antisense, decoy, aptamer and/or 2-5A chimera molecules to one or more targets to alleviate symptoms of a disease.

Administration of Nucleic Acid Molecules

Methods for the delivery of nucleic acid molecules are described in Akhtar *et al.,* 1992, *Trends Cell Bio.*, 2, 139; *Delivery Strategies for Antisense Oligonucleotide Therapeutics*, ed. Akhtar, 1995, Maurer *et al.,* 1999, *Mol. Membr. Biol.,* 16, 129-140; Hofland and Huang,

1999, *Handb. Exp. Pharmacol.*, 137, 165-192; and Lee *et al.*, 2000, *ACS Symp. Ser.*, 752, 184-192, Sullivan *et al.*, PCT WO 94/02595, further describes the general methods for delivery of enzymatic nucleic acid molecules. These protocols can be utilized for the delivery of virtually any nucleic acid molecule. Nucleic acid molecules can be administered to cells by a variety of methods known to those of skill in the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres, or by proteinaceous vectors (O'Hare and Normand, International PCT Publication No. WO 00/53722). Alternatively, the nucleic acid/vehicle combination is locally delivered by direct injection or by use of an infusion pump. Direct injection of the nucleic acid molecules of the invention, whether subcutaneous, intramuscular, or intradermal, can take place using standard needle and syringe methodologies, or by needle-free technologies such as those described in Conry *et al.*, 1999, *Clin. Cancer Res.*, 5, 2330-2337 and Barry *et al.*, International PCT Publication No. WO 99/31262. The molecules of the instant invention can be used as pharmaceutical agents. Pharmaceutical agents prevent, modulate the occurrence, or treat (alleviate a symptom to some extent, preferably all of the symptoms) of a disease state in a patient.

Thus, the invention features a pharmaceutical composition comprising one or more nucleic acid(s) of the invention in an acceptable carrier, such as a stabilizer, buffer, and the like. The negatively charged polynucleotides of the invention can be administered (*e.g.*, RNA, DNA or protein) and introduced into a patient by any standard means, with or without stabilizers, buffers, and the like, to form a pharmaceutical composition. When it is desired to use a liposome delivery mechanism, standard protocols for formation of liposomes can be followed. The compositions of the present invention may also be formulated and used as tablets, capsules or elixirs for oral administration, suppositories for rectal administration, sterile solutions, suspensions for injectable administration, and the other compositions known in the art.

The present invention also includes pharmaceutically acceptable formulations of the compounds described. These formulations include salts of the above compounds, *e.g.*, acid addition salts, for example, salts of hydrochloric, hydrobromic, acetic acid, and benzene sulfonic acid.

A pharmacological composition or formulation refers to a composition or formulation in a form suitable for administration, *e.g.*, systemic administration, into a cell or patient, including for example a human. Suitable forms, in part, depend upon the use or the route of entry, for example oral, transdermal, or by injection. Such forms should not prevent the composition or formulation from reaching a target cell (*i.e.*, a cell to which the negatively

charged nucleic acid is desirable for delivery). For example, pharmacological compositions injected into the blood stream should be soluble. Other factors are known in the art, and include considerations such as toxicity and forms that prevent the composition or formulation from exerting its effect.

By "systemic administration" is meant *in vivo* systemic absorption or accumulation of drugs in the blood stream followed by distribution throughout the entire body. Administration routes which lead to systemic absorption include, without limitation: intravenous, subcutaneous, intraperitoneal, inhalation, oral, intrapulmonary and intramuscular. Each of these administration routes expose the desired negatively charged polymers, *e.g.*, nucleic acids, to an accessible diseased tissue. The rate of entry of a drug into the circulation has been shown to be a function of molecular weight or size. The use of a liposome or other drug carrier comprising the compounds of the instant invention can potentially localize the drug, for example, in certain tissue types, such as the tissues of the reticular endothelial system (RES). A liposome formulation that can facilitate the association of drug with the surface of cells, such as, lymphocytes and macrophages is also useful. This approach may provide enhanced delivery of the drug to target cells by taking advantage of the specificity of macrophage and lymphocyte immune recognition of abnormal cells, such as cancer cells.

By "pharmaceutically acceptable formulation" is meant, a composition or formulation that allows for the effective distribution of the nucleic acid molecules of the instant invention in the physical location most suitable for their desired activity. Nonlimiting examples of agents suitable for formulation with the nucleic acid molecules of the instant invention include: P-glycoprotein inhibitors (such as Pluronic P85), which can enhance entry of drugs into the CNS (Jolliet-Riant and Tillement, 1999, *Fundam. Clin. Pharmacol.*, 13, 16-26); biodegradable polymers, such as poly (DL-lactide-coglycolide) microspheres for sustained release delivery after intracerebral implantation (Emerich, DF *et al*, 1999, *Cell Transplant*, 8, 47-58) (Alkermes, Inc. Cambridge, MA); and loaded nanoparticles, such as those made of polybutylcyanoacrylate, which can deliver drugs across the blood brain barrier and can alter neuronal uptake mechanisms (*Prog Neuropsychopharmacol Biol Psychiatry*, 23, 941-949, 1999). Other non-limiting examples of delivery strategies for the nucleic acid molecules of the instant invention include material described in Boado *et al.*, 1998, *J. Pharm. Sci.*, 87, 1308-1315; Tyler *et al.*, 1999, *FEBS Lett.*, 421, 280-284; Pardridge *et al.*, 1995, *PNAS USA.*, 92, 5592-5596; Boado, 1995, *Adv. Drug Delivery Rev.*, 15, 73-107; Aldrian-Herrada *et al.*, 1998, *Nucleic Acids Res.*, 26, 4910-4916; and Tyler *et al.*, 1999, *PNAS USA.*, 96, 7053-7058.

The invention also features the use of the composition comprising surface-modified liposomes containing poly (ethylene glycol) lipids (PEG-modified, or long-circulating

liposomes or stealth liposomes). These formulations offer a method for increasing the accumulation of drugs in target tissues. This class of drug carriers resists opsonization and elimination by the mononuclear phagocytic system (MPS or RES), thereby enabling longer blood circulation times and enhanced tissue exposure for the encapsulated drug (Lasic *et al. Chem. Rev.* 1995, 95, 2601-2627; Ishiwata *et al., Chem. Pharm. Bull.* 1995, 43, 1005-1011). Such liposomes have been shown to accumulate selectively in tumors, presumably by extravasation and capture in the neovascularized target tissues (Lasic *et al., Science* 1995, 267, 1275-1276; Oku *et al.,*1995, *Biochim. Biophys. Acta*, 1238, 86-90). The long-circulating liposomes enhance the pharmacokinetics and pharmacodynamics of DNA and RNA, particularly compared to conventional cationic liposomes which are known to accumulate in tissues of the MPS (Liu *et al., J. Biol. Chem.* 1995, 42, 24864-24870; Choi *et al.,* International PCT Publication No. WO 96/10391; Ansell *et al.,* International PCT Publication No. WO 96/10390; Holland *et al.,* International PCT Publication No. WO 96/10392). Long-circulating liposomes are also likely to protect drugs from nuclease degradation to a greater extent compared to cationic liposomes, based on their ability to avoid accumulation in metabolically aggressive MPS tissues such as the liver and spleen.

The present invention also includes compositions prepared for storage or administration, which include a pharmaceutically effective amount of the desired compounds in a pharmaceutically acceptable carrier or diluent. Acceptable carriers or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in *Remington's Pharmaceutical Sciences*, Mack Publishing Co. (A.R. Gennaro edit. 1985) hereby incorporated by reference herein. For example, preservatives, stabilizers, dyes and flavoring agents may be provided. These include sodium benzoate, sorbic acid and esters of *p*-hydroxybenzoic acid. In addition, antioxidants and suspending agents may be used.

A pharmaceutically effective dose is that dose required to prevent, inhibit the occurrence of, or treat (alleviate a symptom to some extent, preferably all of the symptoms) a disease state. The pharmaceutically effective dose depends on the type of disease, the composition used, the route of administration, the type of mammal being treated, the physical characteristics of the specific mammal under consideration, concurrent medication, and other factors that those skilled in the medical arts will recognize. Generally, an amount between 0.1 mg/kg and 100 mg/kg body weight/day of active ingredients is administered dependent upon potency of the negatively charged polymer.

The present invention also includes compositions prepared for storage or administration that include a pharmaceutically effective amount of the desired compounds in a pharmaceutically acceptable carrier or diluent. Acceptable carriers or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in *Remington's*

Pharmaceutical Sciences, Mack Publishing Co. (A.R. Gennaro edit. 1985), hereby incorporated by reference herein. For example, preservatives, stabilizers, dyes and flavoring agents can be provided. These include sodium benzoate, sorbic acid and esters of *p*-hydroxybenzoic acid. In addition, antioxidants and suspending agents can be used.

A pharmaceutically effective dose is that dose required to prevent, inhibit the occurrence, or treat (alleviate a symptom to some extent, preferably all of the symptoms) of a disease state. The pharmaceutically effective dose depends on the type of disease, the composition used, the route of administration, the type of mammal being treated, the physical characteristics of the specific mammal under consideration, concurrent medication, and other factors that those skilled in the medical arts will recognize. Generally, an amount between 0.1 mg/kg and 100 mg/kg body weight/day of active ingredients is administered dependent upon potency of the negatively charged polymer.

The nucleic acid molecules of the invention and formulations thereof can be administered orally, topically, parenterally, by inhalation or spray, or rectally in dosage unit formulations containing conventional non-toxic pharmaceutically acceptable carriers, adjuvants and/or vehicles. The term parenteral as used herein includes percutaneous, subcutaneous, intravascular (*e.g.*, intravenous), intramuscular, or intrathecal injection or infusion techniques and the like. In addition, there is provided a pharmaceutical formulation comprising a nucleic acid molecule of the invention and a pharmaceutically acceptable carrier. One or more nucleic acid molecules of the invention can be present in association with one or more non-toxic pharmaceutically acceptable carriers and/or diluents and/or adjuvants, and if desired other active ingredients. The pharmaceutical compositions containing nucleic acid molecules of the invention can be in a form suitable for oral use, for example, as tablets, troches, lozenges, aqueous or oily suspensions, dispersible powders or granules, emulsion, hard or soft capsules, or syrups or elixirs.

Compositions intended for oral use can be prepared according to any method known to the art for the manufacture of pharmaceutical compositions and such compositions can contain one or more such sweetening agents, flavoring agents, coloring agents or preservative agents in order to provide pharmaceutically elegant and palatable preparations. Tablets contain the active ingredient in admixture with non-toxic pharmaceutically acceptable excipients that are suitable for the manufacture of tablets. These excipients can be, for example, inert diluents; such as calcium carbonate, sodium carbonate, lactose, calcium phosphate or sodium phosphate; granulating and disintegrating agents, for example, corn starch, or alginic acid; binding agents, for example starch, gelatin or acacia; and lubricating agents, for example magnesium stearate, stearic acid or talc. The tablets can be uncoated or they can be coated by known techniques. In some cases such coatings can be prepared by

known techniques to delay disintegration and absorption in the gastrointestinal tract and thereby provide a sustained action over a longer period. For example, a time delay material such as glyceryl monosterate or glyceryl distearate can be employed.

Formulations for oral use can also be presented as hard gelatin capsules wherein the active ingredient is mixed with an inert solid diluent, for example, calcium carbonate, calcium phosphate or kaolin, or as soft gelatin capsules wherein the active ingredient is mixed with water or an oil medium, for example peanut oil, liquid paraffin or olive oil.

Aqueous suspensions contain the active materials in admixture with excipients suitable for the manufacture of aqueous suspensions. Such excipients are suspending agents, for example sodium carboxymethylcellulose, methylcellulose, hydropropyl-methylcellulose, sodium alginate, polyvinylpyrrolidone, gum tragacanth and gum acacia; dispersing or wetting agents can be a naturally-occurring phosphatide, for example, lecithin, or condensation products of an alkylene oxide with fatty acids, for example polyoxyethylene stearate, or condensation products of ethylene oxide with long chain aliphatic alcohols, for example heptadecaethyleneoxycetanol, or condensation products of ethylene oxide with partial esters derived from fatty acids and a hexitol such as polyoxyethylene sorbitol monooleate, or condensation products of ethylene oxide with partial esters derived from fatty acids and hexitol anhydrides, for example polyethylene sorbitan monooleate. The aqueous suspensions can also contain one or more preservatives, for example ethyl, or n-propyl p-hydroxybenzoate, one or more coloring agents, one or more flavoring agents, and one or more sweetening agents, such as sucrose or saccharin.

Oily suspensions can be formulated by suspending the active ingredients in a vegetable oil, for example arachis oil, olive oil, sesame oil or coconut oil, or in a mineral oil such as liquid paraffin. The oily suspensions can contain a thickening agent, for example beeswax, hard paraffin or cetyl alcohol. Sweetening agents and flavoring agents can be added to provide palatable oral preparations. These compositions can be preserved by the addition of an anti-oxidant such as ascorbic acid.

Dispersible powders and granules suitable for preparation of an aqueous suspension by the addition of water provide the active ingredient in admixture with a dispersing or wetting agent, suspending agent and one or more preservatives. Suitable dispersing or wetting agents or suspending agents are exemplified by those already mentioned above. Additional excipients, for example sweetening, flavoring and coloring agents, can also be present.

Pharmaceutical compositions of the invention can also be in the form of oil-in-water emulsions. The oily phase can be a vegetable oil or a mineral oil or mixtures of these. Suitable emulsifying agents can be naturally-occurring gums, for example gum acacia or gum

tragacanth, naturally-occurring phosphatides, for example soy bean, lecithin, and esters or partial esters derived from fatty acids and hexitol, anhydrides, for example sorbitan monooleate, and condensation products of the said partial esters with ethylene oxide, for example polyoxyethylene sorbitan monooleate. The emulsions can also contain sweetening and flavoring agents.

Syrups and elixirs can be formulated with sweetening agents, for example glycerol, propylene glycol, sorbitol, glucose or sucrose. Such formulations can also contain a demulcent, a preservative and flavoring and coloring agents. The pharmaceutical compositions can be in the form of a sterile injectable aqueous or oleaginous suspension. This suspension can be formulated according to the known art using those suitable dispersing or wetting agents and suspending agents that have been mentioned above. The sterile injectable preparation can also be a sterile injectable solution or suspension in a non-toxic parentally acceptable diluent or solvent, for example as a solution in 1,3-butanediol. Among the acceptable vehicles and solvents that can be employed are water, Ringer's solution and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose, any bland fixed oil can be employed including synthetic mono-or diglycerides. In addition, fatty acids such as oleic acid find use in the preparation of injectables.

The nucleic acid molecules of the invention can also be administered in the form of suppositories, *e.g.*, for rectal administration of the drug. These compositions can be prepared by mixing the drug with a suitable non-irritating excipient that is solid at ordinary temperatures but liquid at the rectal temperature and will therefore melt in the rectum to release the drug. Such materials include cocoa butter and polyethylene glycols.

Nucleic acid molecules of the invention can be administered parenterally in a sterile medium. The drug, depending on the vehicle and concentration used, can either be suspended or dissolved in the vehicle. Advantageously, adjuvants such as local anesthetics, preservatives and buffering agents can be dissolved in the vehicle.

Dosage levels of the order of from about 0.1 mg to about 140 mg per kilogram of body weight per day are useful in the treatment of the above-indicated conditions (about 0.5 mg to about 7 g per patient per day). The amount of active ingredient that can be combined with the carrier materials to produce a single dosage form varies depending upon the host treated and the particular mode of administration. Dosage unit forms generally contain between from about 1 mg to about 500 mg of an active ingredient.

It is understood that the specific dose level for any particular patient depends upon a variety of factors including the activity of the specific compound employed, the age, body

weight, general health, sex, diet, time of administration, route of administration, and rate of excretion, drug combination and the severity of the particular disease undergoing therapy.

For administration to non-human animals, the composition can also be added to the animal feed or drinking water. It can be convenient to formulate the animal feed and drinking water compositions so that the animal takes in a therapeutically appropriate quantity of the composition along with its diet. It can also be convenient to present the composition as a premix for addition to the feed or drinking water.

The nucleic acid molecules of the present invention may also be administered to a patient in combination with other therapeutic compounds to increase the overall therapeutic effect. The use of multiple compounds to treat an indication may increase the beneficial effects while reducing the presence of side effects.

In one embodiment, the invention compositions suitable for administering nucleic acid molecules of the invention to specific cell types, such as hepatocytes. For example, the asialoglycoprotein receptor (ASGPr) (Wu and Wu, 1987, *J. Biol. Chem.* 262, 4429-4432) is unique to hepatocytes and binds branched galactose-terminal glycoproteins, such as asialoorosomucoid (ASOR). Binding of such glycoproteins or synthetic glycoconjugates to the receptor takes place with an affinity that strongly depends on the degree of branching of the oligosaccharide chain, for example, triatennary structures are bound with greater affinity than biatenarry or monoatennary chains (Baenziger and Fiete, 1980, *Cell,* 22, 611-620; Connolly *et al.*, 1982, *J. Biol. Chem.*, 257, 939-945). Lee and Lee, 1987, *Glycoconjugate J.*, 4, 317-328, obtained this high specificity through the use of N-acetyl-D-galactosamine as the carbohydrate moiety, which has higher affinity for the receptor, compared to galactose. This "clustering effect" has also been described for the binding and uptake of mannosyl-terminating glycoproteins or glycoconjugates (Ponpipom *et al.*, 1981, *J. Med. Chem.*, 24, 1388-1395). The use of galactose and galactosamine based conjugates to transport exogenous compounds across cell membranes can provide a targeted delivery approach to the treatment of liver disease such as HBV infection or hepatocellular carcinoma. The use of bioconjugates can also provide a reduction in the required dose of therapeutic compounds required for treatment. Furthermore, therapeutic bioavialability, pharmacodynamics, and pharmacokinetic parameters can be modulated through the use of nucleic acid bioconjugates of the invention.

Alternatively, certain of the nucleic acid molecules of the instant invention can be expressed within cells from eukaryotic promoters (*e.g.*, Izant and Weintraub, 1985, *Science,* 229, 345; McGarry and Lindquist, 1986, *Proc. Natl. Acad. Sci.*, USA 83, 399; Scanlon *et al.*, 1991, *Proc. Natl. Acad. Sci. USA*, 88, 10591-5; Kashani-Sabet *et al.*, 1992, *Antisense Res. Dev.*, 2, 3-15; Dropulic *et al.*, 1992, *J. Virol.*, 66, 1432-41; Weerasinghe *et al.*, 1991, *J. Virol.*, 65, 5531-4; Ojwang *et al.*, 1992, *Proc. Natl. Acad. Sci. USA*, 89, 10802-6; Chen *et*

al., 1992, *Nucleic Acids Res.*, 20, 4581-9; Sarver *et al.*, 1990 *Science*, 247, 1222-1225; Thompson *et al.*, 1995, *Nucleic Acids Res.*, 23, 2259; Good *et al.*, 1997, *Gene* Therapy, 4, 45; all of these references are hereby incorporated in their totalities by reference herein). Those skilled in the art realize that any nucleic acid can be expressed in eukaryotic cells from the appropriate DNA/RNA vector. The activity of such nucleic acids can be augmented by their release from the primary transcript by a ribozyme (Draper *et al.*, PCT WO 93/23569, and Sullivan *et al.*, PCT WO 94/02595; Ohkawa *et al.*, 1992, *Nucleic Acids Symp. Ser.*, 27, 15-6; Taira *et al.*, 1991, *Nucleic Acids Res.*, 19, 5125-30; Ventura *et al.*, 1993, *Nucleic Acids Res.*, 21, 3249-55; Chowrira *et al.*, 1994, *J. Biol. Chem.*, 269, 25856; all of these references are hereby incorporated in their totality by reference herein).

In another aspect of the invention, RNA molecules of the present invention are preferably expressed from transcription units (see, for example, Couture *et al.*, 1996, *TIG.*, 12, 510) inserted into DNA or RNA vectors. The recombinant vectors are preferably DNA plasmids or viral vectors. Ribozyme expressing viral vectors could be constructed based on, but not limited to, adeno-associated virus, retrovirus, adenovirus, or alphavirus. Preferably, the recombinant vectors capable of expressing the nucleic acid molecules are delivered as described above, and persist in target cells. Alternatively, viral vectors may be used that provide for transient expression of nucleic acid molecules. Such vectors might be repeatedly administered as necessary. Once expressed, the nucleic acid molecule binds to the target mRNA. Delivery of nucleic acid molecule expressing vectors could be systemic, such as by intravenous or intra-muscular administration, by administration to target cells ex-planted from the patient followed by reintroduction into the patient, or by any other means that would allow for introduction into the desired target cell (for a review see Couture *et al.*, 1996, *TIG.*, 12, 510).

In one aspect, the invention features an expression vector comprising a nucleic acid sequence encoding at least one of the nucleic acid molecules of the instant invention is disclosed. The nucleic acid sequence encoding the nucleic acid molecule of the instant invention is operable linked in a manner which allows expression of that nucleic acid molecule.

In another aspect the invention features an expression vector comprising: a) a transcription initiation region (*e.g.*, eukaryotic pol I, II or III initiation region); b) a transcription termination region (*e.g.*, eukaryotic pol I, II or III termination region); c) a nucleic acid sequence encoding at least one of the nucleic acid catalyst of the instant invention; and wherein said sequence is operably linked to said initiation region and said termination region, in a manner which allows expression and/or delivery of said nucleic acid molecule. The vector may optionally include an open reading frame (ORF) for a protein

operably linked on the 5' side or the 3'-side of the sequence encoding the nucleic acid catalyst of the invention; and/or an intron (intervening sequences).

Transcription of the nucleic acid molecule sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, etc.) present nearby. Prokaryotic RNA polymerase promoters are also used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990, *Proc. Natl. Acad. Sci. U S A*, 87, 6743-7; Gao and Huang 1993, *Nucleic Acids Res.*, 21, 2867-72; Lieber et al., 1993, *Methods Enzymol.*, 217, 47-66; Zhou et al., 1990, *Mol. Cell. Biol.*, 10, 4529-37). All of these references are incorporated by reference herein. Several investigators have demonstrated that nucleic acid molecules, such as ribozymes expressed from such promoters can function in mammalian cells (e.g. Kashani-Sabet et al., 1992, *Antisense Res. Dev.*, 2, 3-15; Ojwang et al., 1992, *Proc. Natl. Acad. Sci. U S A*, 89, 10802-6; Chen *et al.*, 1992, *Nucleic Acids Res.*, 20, 4581-9; Yu et al., 1993, *Proc. Natl. Acad. Sci. U S A*, 90, 6340-4; L'Huillier *et al.*, 1992, *EMBO J.*, 11, 4411-8; Lisziewicz *et al.*, 1993, *Proc. Natl. Acad. Sci. U. S. A*, 90, 8000-4; Thompson *et al.*, 1995, *Nucleic Acids Res.*, 23, 2259; Sullenger & Cech, 1993, *Science*, 262, 1566). More specifically, transcription units such as the ones derived from genes encoding U6 small nuclear (snRNA), transfer RNA (tRNA) and adenovirus VA RNA are useful in generating high concentrations of desired RNA molecules such as ribozymes in cells (Thompson *et al., supra*; Couture and Stinchcomb, *1996, supra*; Noonberg *et al.*, 1994, *Nucleic Acid Res.*, 22, 2830; Noonberg *et al.*, US Patent No. 5,624,803; Good *et al.*, 1997, *Gene Ther.*, 4, 45; Beigelman *et al.*, International PCT Publication No. *WO 96/18736*; all of these publications are incorporated by reference herein). The above ribozyme transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated virus vectors), or viral RNA vectors (such as retroviral or alphavirus vectors) (for a review see Couture and Stinchcomb, *1996, supra)*.

In yet another aspect, the invention features an expression vector comprising nucleic acid sequence encoding at least one of the nucleic acid molecules of the invention, in a manner that allows expression of that nucleic acid molecule. The expression vector comprises in one embodiment; a) a transcription initiation region; b) a transcription termination region; c) a nucleic acid sequence encoding at least one said nucleic acid molecule; and wherein said sequence is operably linked to said initiation region and said termination region, in a manner which allows expression and/or delivery of said nucleic acid molecule. In another embodiment, the expression vector comprises: a) a transcription initiation region; b) a

transcription termination region; c) an open reading frame; d) a nucleic acid sequence encoding at least one said nucleic acid molecule, wherein said sequence is operably linked to the 3'-end of said open reading frame; and wherein said sequence is operably linked to said initiation region, said open reading frame and said termination region, in a manner which allows expression and/or delivery of said nucleic acid molecule. In yet another embodiment, the expression vector comprises: a) a transcription initiation region; b) a transcription termination region; c) an intron; d) a nucleic acid sequence encoding at least one said nucleic acid molecule; and wherein said sequence is operably linked to said initiation region, said intron and said termination region, in a manner which allows expression and/or delivery of said nucleic acid molecule. In another embodiment, the expression vector comprises: a) a transcription initiation region; b) a transcription termination region; c) an intron; d) an open reading frame; e) a nucleic acid sequence encoding at least one said nucleic acid molecule, wherein said sequence is operably linked to the 3'-end of said open reading frame; and wherein said sequence is operably linked to said initiation region, said intron, said open reading frame and said termination region, in a manner which allows expression and/or delivery of said nucleic acid molecule.

Interferons

Type I interferons (IFN) are a class of natural cytokines that includes a family of greater than 25 IFN-α (Pesta, 1986, *Methods Enzymol.* 119, 3-14) as well as IFN-β, and IFN-ω. Although evolutionarily derived from the same gene (Diaz *et al.*, 1994, *Genomics* 22, 540-552), there are many differences in the primary sequence of these molecules, implying an evolutionary divergence in biologic activity. All type I IFN share a common pattern of biologic effects that begin with binding of the IFN to the cell surface receptor (Pfeffer & Strulovici, 1992, Transmembrane secondary messengers for IFN-α/β. In: *Interferon. Principles and Medical Applications.*, S. Baron, D.H. Coopenhaver, F. Dianzani, W.R. Fleischmann Jr., T.K. Hughes Jr., G.R. Kimpel, D.W. Niesel, G.J. Stanton, and S.K. Tyring, eds. 151-160). Binding is followed by activation of tyrosine kinases, including the Janus tyrosine kinases and the STAT proteins, which leads to the production of several IFN-stimulated gene products (Johnson *et al.*, 1994, *Sci. Am.* 270, 68-75). The IFN-stimulated gene products are responsible for the pleotropic biologic effects of type I IFN, including antiviral, antiproliferative, and immunomodulatory effects, cytokine induction, and HLA class I and class II regulation (Pestka *et al.*, 1987, *Annu. Rev. Biochem* 56, 727). Examples of IFN-stimulated gene products include 2-5-oligoadenylate synthetase (2-5 OAS), β_2-microglobulin, neopterin, p68 kinases, and the Mx protein (Chebath & Revel, 1992, The 2-5 A system: 2-5 A synthetase, isospecies and functions. In: *Interferon. Principles and Medical Applications.* S. Baron, D.H. Coopenhaver, F. Dianzani, W.R. Jr. Fleischmann, T.K. Jr Hughes, G.R. Kimpel, D.W. Niesel, G.J. Stanton, and S.K. Tyring, eds., pp. 225-236;

Samuel, 1992, The RNA-dependent P1/eIF-2α protein kinase. In: *Interferon. Principles and Medical Applications*. S. Baron, D.H. Coopenhaver, F. Dianzani, W.R. Fleischmann Jr., T.K. Hughes Jr., G.R. Kimpel, D.W. Niesel, G.H. Stanton, and S.K. Tyring, eds. 237-250; Horisberger, 1992, MX protein: function and Mechanism of Action. In: *Interferon. Principles and Medical Applications*. S. Baron, D.H. Coopenhaver, F. Dianzani, W.R. Fleischmann Jr., T.K. Hughes Jr., G.R. Kimpel, D.W. Niesel, G.H. Stanton, and S.K. Tyring, eds. 215-224). Although all type I IFN have similar biologic effects, not all the activities are shared by each type I IFN, and, in many cases, the extent of activity varies quite substantially for each IFN subtype (Fish *et al*, 1989, *J. Interferon Res.* 9, 97-114; Ozes *et al.*, 1992, *J. Interferon Res.* 12, 55-59). More specifically, investigations into the properties of different subtypes of IFN-α and molecular hybrids of IFN-α have shown differences in pharmacologic properties (Rubinstein, 1987, *J. Interferon Res.* 7, 545-551). These pharmacologic differences can arise from as few as three amino acid residue changes (Lee *et al.*, 1982, *Cancer Res.* 42, 1312-1316).

Eighty-five to 166 amino acids are conserved in the known IFN-α subtypes. Excluding the IFN-α pseudogenes, there are approximately 25 known distinct IFN-α subtypes. Pairwise comparisons of these nonallelic subtypes show primary sequence differences ranging from 2% to 23%. In addition to the naturally occurring IFNs, a non-natural recombinant type I interferon known as consensus interferon (CIFN) has been synthesized as a therapeutic compound (Tong *et al.*, 1997, *Hepatology* 26, 747-754).

Interferon is currently in use for at least 12 different indications including infectious and autoimmune diseases and cancer (Borden, 1992, *N. Engl. J. Med.* 326, 1491-1492). For autoimmune diseases IFN has been utilized for treatment of rheumatoid arthritis, multiple sclerosis, and Crohn's disease. For treatment of cancer IFN has been used alone or in combination with a number of different compounds. Specific types of cancers for which IFN has been used include squamous cell carcinomas, melanomas, hypernephromas, hemangiomas, hairy cell leukemia, and Kaposi's sarcoma. In the treatment of infectious diseases, IFNs increase the phagocytic activity of macrophages and cytotoxicity of lymphocytes and inhibits the propagation of cellular pathogens. Specific indications for which IFN has been used as treatment include: hepatitis B, human papillomavirus types 6 and 11 (i.e. genital warts) (Leventhal *et al.*, 1991, *N Engl J Med* 325, 613-617), chronic granulomatous disease, and hepatitis C virus.

Numerous well controlled clinical trials using IFN-alpha in the treatment of chronic HCV infection have demonstrated that treatment three times a week results in lowering of serum ALT values in approximately 50% (range 40% to 70%) of patients by the end of 6 months of therapy (Davis *et al.*, 1989, The new England Journal of Medicine 321, 1501-

1506; Marcellin et al., 1991, *Hepatology* 13, 393-397; Tong *et al.*, 1997, Hepatology 26, 747-754; Tong et al., *Hepatology* 26, 1640-1645). However, following cessation of interferon treatment, approximately 50% of the responding patients relapsed, resulting in a "durable" response rate as assessed by normalization of serum ALT concentrations of approximately 20 to 25%. In addition, studies that have examined six months of type 1 interferon therapy using changes in HCV RNA values as a clinical endpoint have demonstrated that up to 35% of patients will have a loss of HCV RNA by the end of therapy (Tong *et al.*, 1997, supra). However, as with the ALT endpoint, about 50% of the patients relapse six months following cessation of therapy resulting in a durable virologic response of only 12% (23). Studies that have examined 48 weeks of therapy have demonstrated that the sustained virological response is up to 25%.

Pegylated interferons, ie. interferons conjugated with polyethylene glycol (PEG), have demonstrated improved characteristics over interferon. Advantages incurred by PEG conjugation can include an improved pharmacokinetic profile compared to interferons lacking PEG, thus imparting more convenient dosing regimes, improved tolerance, and improved antiviral efficacy. Such improvements have been demonstrated in clinical studies of both polyethylene glycol interferon alfa-2a (PEGASYS, Roche) and polyethylene glycol interferon alfa-2b (VIRAFERON PEG, PEG-INTRON, Enzon/Schering Plough).

Enzymatic nucleic acid molecules in combination with interferons and polyethylene glycol interferons have the potential to improve the effectiveness of treatment of HCV or any of the other indications discussed above. Enzymatic nucleic acid molecules targeting RNAs associated with diseases such as infectious diseases, autoimmune diseases, and cancer, can be used individually or in combination with other therapies such as interferons and polyethylene glycol interferons and to achieve enhanced efficacy.

Examples:

The following are non-limiting examples showing the selection, isolation, synthesis and activity of nucleic acids of the instant invention. These examples demonstrate the selection and design of Antisense, Hammerhead, DNAzyme, NCH, Amberzyme, Zinzyme or G-Cleaver ribozyme molecules and binding/cleavage sites within HBV and HCV RNA. The following examples also demonstrate the selection and design of nucleic acid decoy molecules that target HBV reverse transcriptase. The following examples also demonstrate the use of enzymatic nucleic acid molecules that cleave HCV RNA. The methods described herein represent a scheme by which nucleic acid molecules can be derived that cleave other RNA targets required for HCV replication.

Example 1: Identification of Potential Target Sites in Human HBV RNA

The sequence of human HBV was screened for accessible sites using a computer-folding algorithm. Regions of the RNA that did not form secondary folding structures and contained potential ribozyme and/or antisense binding/cleavage sites were identified. The sequences of these cleavage sites are shown in **Tables IV - XI.**

Example 2: Selection of Enzymatic Nucleic Acid Cleavage Sites in Human HBV RNA

Ribozyme target sites were chosen by analyzing sequences of Human HBV (accession number: AF100308.1) and prioritizing the sites on the basis of folding. Ribozymes were designed that could bind each target and were individually analyzed by computer folding (Christoffersen *et al.,* 1994 *J. Mol. Struc. Theochem,* 311, 273; Jaeger *et al.*, 1989, *Proc. Natl. Acad. Sci. USA*, **86**, 7706) to assess whether the ribozyme sequences fold into the appropriate secondary structure. Those ribozymes with unfavorable intramolecular interactions between the binding arms and the catalytic core were eliminated from consideration. As noted herein, varying binding arm lengths can be chosen to optimize activity. Generally, at least 5 bases on each arm are able to bind to, or otherwise interact with, the target RNA.

Example 3: Chemical Synthesis and Purification of Ribozymes and Antisense for Efficient Cleavage and/or blocking of HBV RNA

Ribozymes and antisense constructs were designed to anneal to various sites in the RNA message. The binding arms of the ribozymes are complementary to the target site sequences described above, while the antisense constructs are fully complementary to the target site sequences described above. The ribozymes and antisense constructs were chemically synthesized. The method of synthesis used followed the procedure for normal RNA synthesis as described above and in Usman et al., (1987 *J. Am. Chem. Soc.*, 109, 7845), Scaringe et al., (1990 *Nucleic Acids Res.*, 18, 5433) and Wincott et al., *supra*, and made use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. The average stepwise coupling yields were typically >98%.

Ribozymes and antisense constructs were also synthesized from DNA templates using bacteriophage T7 RNA polymerase (Milligan and Uhlenbeck, 1989, *Methods Enzymol.* 180, 51). Ribozymes and antisense constructs were purified by gel electrophoresis using general methods or were purified by high pressure liquid chromatography (HPLC; see Wincott et al., *supra*; the totality of which is hereby incorporated herein by reference) and were resuspended in water. The sequences of the chemically synthesized ribozymes used in this study are shown below in **Table XI**.

Example 4: Ribozyme Cleavage of HBV RNA Target *in vitro*

Ribozymes targeted to the human HBV RNA are designed and synthesized as described above. These ribozymes can be tested for cleavage activity *in vitro*, for example using the following procedure. The target sequences and the nucleotide location within the HBV RNA are given in **Tables IV-XI**.

Cleavage Reactions: Full-length or partially full-length, internally-labeled target RNA for ribozyme cleavage assay is prepared by *in vitro* transcription in the presence of [α-^{32}p] CTP, passed over a G 50 Sephadex® column by spin chromatography and used as substrate RNA without further purification. Alternately, substrates are 5'-^{32}P-end labeled using T4 polynucleotide kinase enzyme. Assays are performed by pre-warming a 2X concentration of purified ribozyme in ribozyme cleavage buffer (50 mM Tris-HCl, pH 7.5 at 37°C, 10 mM $MgCl_2$) and the cleavage reaction was initiated by adding the 2X ribozyme mix to an equal volume of substrate RNA (maximum of 1-5 nM) that was also pre-warmed in cleavage buffer. As an initial screen, assays are carried out for 1 hour at 37°C using a final concentration of either 40 nM or 1 mM ribozyme, *i.e.*, ribozyme excess. The reaction is quenched by the addition of an equal volume of 95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol after which the sample is heated to 95°C for 2 minutes, quick chilled and loaded onto a denaturing polyacrylamide gel. Substrate RNA and the specific RNA cleavage products generated by ribozyme cleavage are visualized on an autoradiograph of the gel. The percentage of cleavage is determined by Phosphor Imager® quantitation of bands representing the intact substrate and the cleavage products.

Example 5: Transfection of HepG2 Cells with psHBV-1 and Ribozymes

The human hepatocellular carcinoma cell line Hep G2 was grown in Dulbecco's modified Eagle media supplemented with 10% fetal calf serum, 2 mM glutamine, 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, 25 mM Hepes, 100 units penicillin, and 100 μg/ml streptomycin. To generate a replication competent cDNA, prior to transfection the HBV genomic sequences are excised from the bacterial plasmid sequence contained in the psHBV-1 vector (Those skilled in the art understand that other methods may be used to generate a replication competent cDNA). This was done with an EcoRI and Hind III restriction digest. Following completion of the digest, a ligation was performed under dilute conditions (20 μg/ml) to favor intermolecular ligation. The total ligation mixture was then concentrated using Qiagen spin columns.

Secreted alkaline phosphatase (SEAP) was used to normalize the HBsAg levels to control for transfection variability. The pSEAP2-TK control vector was constructed by ligating a Bgl II-Hind III fragment of the pRL-TK vector (Promega), containing the herpes

simplex virus thymidine kinase promoter region, into *Bgl* II/*Hind* III digested pSEAP2-Basic (Clontech). Hep G2 cells were plated (3 x 10^4 cells/well) in 96-well microtiter plates and incubated overnight. A lipid/DNA/ribozyme complex was formed containing (at final concentrations) cationic lipid (15 μg/ml), prepared psHBV-1 (4.5 μg/ml), pSEAP2-TK (0.5 μg/ml), and ribozyme (100 μM). Following a 15 min. incubation at 37° C, the complexes were added to the plated Hep G2 cells. Media was removed from the cells 96 hr. post-transfection for HBsAg and SEAP analysis.

Transfection of the human hepatocellular carcinoma cell line, Hep G2, with replication competent HBV DNA results in the expression of HBV proteins and the production of virions. To investigate the potential use of ribozymes for the treatment of chronic HBV infection, a series of ribozymes that target the 3' terminus of the HBV genome have been synthesized. Ribozymes targeting this region have the potential to cleave all four major HBV RNA transcripts as well as the potential to block the production of HBV DNA by cleavage of the pregenomic RNA. To test the efficacy of these HBV ribozymes, they were co-transfected with HBV genomic DNA into Hep G2 cells, and the subsequent levels of secreted HBV surface antigen (HBsAg) were analyzed by ELISA. To control for variability in transfection efficiency, a control vector which expresses secreted alkaline phosphatase (SEAP), was also co-transfected. The efficacy of the HBV ribozymes was determined by comparing the ratio of HBsAg:SEAP and/or HBeAg:SEAP to that of a scrambled attenuated control (SAC) ribozyme. Twenty-five ribozymes (RPI18341, RPI18356, RPI18363, RPI18364, RPI18365, RPI18366, RPI18367, RPI18368, RPI18369, RPI18370, RPI18371, RPI18372, RPI18373, RPI18374, RPI18303, RPI18405, RPI18406, RPI18407, RPI18408, RPI18409, RPI18410, RPI18411, RPI18418, RPI18419, and RPI18422) have been identified which cause a reduction in the levels of HBsAg and/or HBeAg as compared to the corresponding SAC ribozyme. In addition, loop variant anti-HBV ribozymes targeting site 273 were tested using this system, the results of this study are summarized in **Figure 10**. As indicated in the figure, the ribozymes tested demonstrate significant reduction in HepG2 HBsAg levels as compared to a scrambled attenuated core ribozyme control, with RPI 22650 and RPI 22649 showing the greatest decrease in HBsAg levels.

Example 6: Analysis of HBsAg and SEAP Levels Following Ribozyme Treatment

Immulon 4 (Dynax) microtiter wells were coated overnight at 4° C with anti-HBsAg Mab (Biostride B88-95-31ad,ay) at 1 μg/ml in Carbonate Buffer (Na2CO3 15 mM, NaHCO3 35 mM, pH 9.5). The wells were then washed 4x with PBST (PBS, 0.05% Tween® 20) and blocked for 1 hr at 37° C with PBST, 1% BSA. Following washing as above, the wells were dried at 37° C for 30 min. Biotinylated goat ant-HBsAg (Accurate YVS1807) was diluted 1:1000 in PBST and incubated in the wells for 1 hr. at 37° C. The wells were washed 4x with

PBST. Streptavidin/Alkaline Phosphatase Conjugate (Pierce 21324) was diluted to 250 ng/ml in PBST, and incubated in the wells for 1 hr. at 37° C. After washing as above, p-nitrophenyl phosphate substrate (Pierce 37620) was added to the wells, which were then incubated for 1 hr. at 37° C. The optical density at 405 nm was then determined. SEAP levels were assayed using the Great EscAPe® Detection Kit (Clontech K2041-1), as per the manufacturers instructions.

Example 7: X-gene Reporter Assay

The effect of ribozyme treatment on the level of transactivation of a SV40 promoter driven firefly luciferase gene by the HBV X-protein was analyzed in transfected Hep G2 cells. As a control for variability in transfection efficiency, a Renilla luciferase reporter driven by the TK promoter, which is not transactivated by the X protein, was used. Hep G2 cells were plated (3×10^4 cells/well) in 96-well microtiter plates and incubated overnight. A lipid/DNA/ribozyme complex was formed containing (at final concentrations) cationic lipid (2.4 μg/ml), the X-gene vector pSBDR(2.5 μg/ml), the firefly reporter pSV40HCVluc (0.5 μg/ml), the Renilla luciferase control vector pRL-TK (0.5 μg/ml), and ribozyme (100 μM). Following a 15 min. incubation at 37° C, the complexes were added to the plated Hep G2 cells. Levels of firefly and Renilla luciferase were analyzed 48 hr. post transfection, using Promega's Dual-Luciferase Assay System.

The HBV X protein is a transactivator of a number of viral and cellular genes. Ribozymes which target the X region were tested for their ability to cause a reduction in X protein transactivation of a firefly luciferase gene driven by the SV40 promoter in transfected Hep G2 cells. As a control for transfection variability, a vector containing the Renilla luciferase gene driven by the TK promotor, which is not activated by the X protein, was included in the co-transfections. The efficacy of the HBV ribozymes was determined by comparing the ratio of firefly luciferase: Renilla luciferase to that of a scrambled attenuated control (SAC) ribozyme. Eleven ribozymes (RPI18365, RPI18367, RPI18368, RPI18371, RPI18372, RPI18373, RPI18405, RPI18406, RPI18411, RPI18418, RPI18423) were identified which cause a reduction in the level of transactivation of a reporter gene by the X protein, as compared to the corresponding SAC ribozyme.

Example 8: HBV transgenic mouse study A

A transgenic mouse strain (founder strain 1.3.32 with a C57B1/6 background) that expresses HBV RNA and forms HBV viremia (Morrey *et al.*, 1999, *Antiviral Res.*, 42, 97-108; Guidotti *et al.*, 1995, *J. Virology*, 69, 10, 6158-6169) was utilized to study the *in vivo* activity of ribozymes (RPI.18341, RPI.18371, RPI.18372, and RPI.18418) of the instant invention. This model is predictive in screening for anti-HBV agents. Ribozyme or the

equivalent volume of saline was administered via a continuous s.c. infusion using Alzet® mini-osmotic pumps for 14 days. Alzet® pumps were filled with test material(s) in a sterile fashion according to the manufacturer's instructions. Prior to *in vivo* implantation, pumps were incubated at 37°C overnight (≥ 18 hours) to prime the flow modulators. On the day of surgery, animals were lightly anesthetized with a ketamine/xylazine cocktail (94 mg/kg and 6 mg/kg, respectively; 0.3 ml, IP). Baseline blood samples (200 µl) were obtained from each animal *via* a retro-orbital bleed. For animals in groups 1-5 (**Table XII**), a 2 cm area near the base of the tail was shaved and cleansed with betadine surgical scrub and sequentially with 70% alcohol. A 1 cm incision in the skin was made with a #15 scalpel blade or a blunt pair of scissors near the base of the tail. Forceps were used to open a pocket rostrally (*ie.*, towards the head) by spreading apart the subcutaneous connective tissue. The pump was inserted with the delivery portal pointing away from the incision. Wounds were closed with sterile 9-mm stainless steel clips or with sterile 4-0 suture. Animals were then allowed to recover from anesthesia on a warm heating pad before being returned to their cage. Wounds were checked daily. Clips or sutures were replaced as needed. Incisions typically healed completely within 7 days post-op. Animals were then deeply anesthetized with the ketamine/xylazine cocktail (150 mg/kg and 10 mg/kg, respectively; 0.5 ml, IP) on day 14 post pump implantation. A midline thoracotomy/ laparatomy was performed to expose the abdominal cavity and the thoracic cavity. The left ventricle was cannulated at the base and animals exsanguinated using a 23G needle and 1 ml syringe. Serum was separated, frozen and analyzed for HBV DNA and antigen levels. Experimental groups were compared to the saline control group in respect to percent change from day 0 to day 14. HBV DNA was assayed by quantitative PCR.

Results

Table XII is a summary of the group designation and dosage levels used in this HBV transgenic mouse study. Baseline blood samples were obtained *via* a retroorbital bleed and animals (N=10/group) received anti-HBV ribozymes (100 mg/kg/day) as a continuous SC infusion. After 14 days, animals treated with a ribozyme targeting site 273 (RPI.18341) of the HBV RNA showed a significant reduction in serum HBV DNA concentration, compared to the saline treated animals as measured by a quantitative PCR assay. More specifically, the saline treated animals had a 69% increase in serum HBV DNA concentrations over this 2-week period while treatment with the 273 ribozyme (RPI.18341) resulted in a 60% decrease in serum HBV DNA concentrations. Ribozymes directed against sites 1833 (RPI.18371), 1873 (RPI.18418), and 1874 (RPI.18372) decreased serum HBV DNA concentrations by 49%, 15% and 16%, respectively.

Example 9: HBV transgenic mouse study B

A transgenic mouse strain (founder strain 1.3.32 with a C57B1/6 background) that expresses HBV RNA and forms HBV viremia (Morrey *et al.*, 1999, *Antiviral Res.*, 42, 97-108; Guidotti *et al.*, 1995, *J. Virology*, 69, 10, 6158-6169) was utilized to study the *in vivo* activity of ribozymes (RPI.18341 and RPI.18371) of the instant invention. This model is predictive in screening for anti-HBV agents. Ribozyme or the equivalent volume of saline was administered via a continuous s.c. infusion using Alzet® mini-osmotic pumps for 14 days. Alzet® pumps were filled with test material(s) in a sterile fashion according to the manufacturer's instructions. Prior to *in vivo* implantation, pumps were incubated at 37°C overnight (≥ 18 hours) to prime the flow modulators. On the day of surgery, animals were lightly anesthetized with a ketamine/xylazine cocktail (94 mg/kg and 6 mg/kg, respectively; 0.3 ml, IP). Baseline blood samples (200 µl) were obtained from each animal *via* a retro-orbital bleed. For animals in groups 1-10 (**Table XIII**), a 2 cm area near the base of the tail was shaved and cleansed with betadine surgical scrub and sequentially with 70% alcohol. A 1 cm incision in the skin was made with a #15 scalpel blade or a blunt pair of scissors near the base of the tail. Forceps were used to open a pocket rostrally (*ie.*, towards the head) by spreading apart the subcutaneous connective tissue. The pump was inserted with the delivery portal pointing away from the incision. Wounds were closed with sterile 9-mm stainless steel clips or with sterile 4-0 suture. Animals were then allowed to recover from anesthesia on a warm heating pad before being returned to their cage. Wounds were checked daily. Clips or sutures were replaced as needed. Incisions typically healed completely within 7 days post-op. Animals were then deeply anesthetized with the ketamine/xylazine cocktail (150 mg/kg and 10 mg/kg, respectively; 0.5 ml, IP) on day 14 post pump implantation. A midline thoracotomy/ laparatomy was performed to expose the abdominal cavity and the thoracic cavity. The left ventricle was cannulated at the base and animals exsanguinated using a 23G needle and 1 ml syringe. Serum was separated, frozen and analyzed for HBV DNA and antigen levels. Experimental groups were compared to the saline control group in respect to percent change from day 0 to day 14. HBV DNA was assayed by quantitative PCR. Additionally, mice treated with 3TC® by oral gavage at a dose of 300 mg/kg/day for 14 days (group 11, **Table XIII**) were used as a positive control.

Results

Table XIII is a summary of the group designation and dosage levels used in this HBV transgenic mouse study. Baseline blood samples were obtained *via* a retroorbital bleed and animals (N=15/group) received anti-HBV ribozymes (100 mg/kg/day, 30 mg/kg/day, 10 mg/kg/day) as a continuous SC infusion. The results of this study are summarized in **Figures 6, 7, and 8**. As **Figures 6, 7,** and **8** demonstrate, Ribozymes directed against sites 273 (RPI.18341) and 1833 (RPI.18371) demonstrate reduction in the serum HBV DNA levels following 14 days of ribozyme treatment in HBV transgenic mice, as compared to scrambled attenuated core (SAC) ribozyme and saline controls. Furthermore, these ribozymes provide similar, and in some cases, greater reduction of serum HBV DNA levels, as compared to the 3TC® positive control, at lower doses than the 3TC® positive control.

Example 10: HBV DNA reduction in HepG2.2.15 cells

Ribozyme treatment of HepG2.2.15 cells was performed in a 96-well plate format, with 12 wells for each different ribozyme tested (RPI.18341, RPI.18371, RPI.18372, RPI.18418, RPI.20599SAC). HBV DNA levels in the media collected between 120 and 144 hours following transfection was determined using the Roche Amplicor HBV Assay. Treatment with RPI.18341 targeting site 273 resulted in a significant ($P<0.05$) decrease in HBV DNA levels of 62% compared to the SAC (RPI.20599). Treatment with RPI.18371 (site 1833) or RPI.18372 (site 1874) resulted in reductions in HBV DNA levels of 55% and 58% respectively, as compared to treatment with the SAC RPI.20599 (see **Figure 9**).

Example 11: RPI 18341 combination treatment with Lamivudine/Infergen®

The therapeutic use of nucleic acid molecules of the invention either alone or in combination with current therapies, for example lamivudine or type 1 IFN, can lead to improved HBV treatment modalities. To assess the potential of combination therapy, HepG2 cells transfected with a replication competent HBV cDNA, were treated with RPI 18341(HepBzyme™), Infergen® (Amgen, Thousand Oaks Ca), and/or Lamivudine (Epivir®: GlaxoSmithKline, Research Triangle Park NC) either alone or in combination. Results indicated that combination treatment with either RPI 18341 plus Infergen® or combination of RPI 18341 plus lamivudine results in additive down regulation of HBsAg expression ($P<0.001$). These studies can be applied to the treatment of lamivudine resistant cells to further assses the potential for combination therapy of RPI 18341 plus currently available therapies for the treatment of chronic Hepatitis B.

Hep G2 cells were plated (2 x 104 cells/well) in 96-well microtiter plates and incubated overnight. A cationic lipid/DNA/ribozyme complex was formed containing (at final

concentrations) lipid (11-15 μg/mL), re-ligated psHBV-1 (4.5 μg/mL) and ribozyme (100-200 nM) in growth media. Following a 15 min incubation at 37°C, 20 μL of the complex was added to the plated Hep G2 cells in 80 μL of growth media minus antibiotics. For combination treatment with interferon, interferon (Infergen®, Amgen, Thousand Oaks CA) was added at 24 hr post-transfection and then incubated for an additional 96 hr. In the case of co-treatment with Lamivudine (3TC®), the ribozyme-containing cell culture media was removed at 120 hr post-transfection, fresh media containing Lamivudine (Epivir®: GlaxoSmithKline, Research Triangle Park NC) was added, and then incubated for an additional 48 hours. Treatment with Lamivudine or interferon individually was done on Hep G2 cells transfected with the pSHBV-1 vector alone and then treated identically to the co-treated cells. All transfections were performed in triplicate. Analysis of HBsAg levels was performed using the Diasorin HBsAg ELISA kit.

Results

At either 500 or 1000 units of Infergen®, the addition of 200 nM of RPI.18341 results in a 75-77% increase in anti-HBV activity as judged by the level of HBsAg secreted from the treated Hep G2 cells. Conversely, the anti-HBV activity of RPI.18341(at 200 nM) is increased 31-39% when used in combination of 500 or 1000 units of Infergen® (**Figure 11**).

At 25 nM Lamivudine (3TC®), the addition of 100 nM of RPI.18341 results in a 48% increase in anti-HBV activity as judged by the level of HBsAg secreted from treated Hep G2 cells. Conversely, the anti-HBV activity of RPI.18341 (at 100 nM) is increased 31% when used in combination with 25 nM Lamivudine (**Figure 12**).

Example 13: Modulation of HBV reverse transcriptase

The HBV reverse transcriptase (pol) binds to the 5' stem-loop structure in the HBV pregenomic RNA and synthesizes a four-nucleotide primer from the template UUCA. The reverse transcriptase then translocates to the 3' end of the pregenomic RNA where the primer binds to the UUCA sequence within the DR1 element and begins first-strand synthesis of HBV DNA. A number of short oligos, ranging in size from 4 to 16-mers, were designed to act as competitive inhibitors of the HBV reverse transcriptase primer, either by blocking the primer binding sites on the HBV RNA or by acting as a decoy.

The oligonucleotides and controls were synthesized in all 2'-O-methyl and 2'-O-allyl versions (**Table XV**). The inverse sequence of all oligos were generated to serve as controls. Primary screening of the competitive inhibitors was completed in the HBsAg transfection/ELISA system, in which the oligo is co-transfeceted with a HBV cDNA vector into Hep G2 cells. Following 4 days of incubation, the levels of HBsAg secreted into the cell

culture media were determined by ELISA. Screening of the 2'-O-allyl versions revealed that two of the decoy oligos (RPI.24944 and RPI.24945), consisting of 3x or 4x repeats of the RT primer binding site UUCA, along with the matched inverse controls, displayed considerable activity by decreasing HBsAg levels (**Figure 15**). This dramatic decrease in HBsAg levels is not due to cellular toxicity, because a MTS assay showed no difference in proliferation between any of the treated cells. A follow up experiment with a 5x UUCA repeat, the inverse sequence control, and a matched scrambled control, showed that all three oligos decreased HBsAg levels without cellular toxicity. Screening of the 2'-O-methyl versions of the oligos showed no activity from the 3x and 4x UUCA repeat (**Figure 16**), also suggesting that the anti-HBV effect is perhaps related to the 2'-O-allyl chemistry rather than to sequence specificity.

Screening of the 2'-O-methyl oligos did show that the 2'-O-methyl 2x UUCA repeat, RPI.24986, displayed activity in decreasing HBsAg levels as compared to the inverse control, RPI.24950. A dose response experiment showed that at the lower concentrations of 100 and 200 nM, RPI.24986 showed greater activity in decreasing HbsAg levels as compared to the inverse control RPI.24950 (**Figure 17**).

Example 14: Modulation of HBV transcription via Oligonucleotides targeting the Enchancer I core region of HBV DNA

In an effort to block HBV replication, oligonucleotides were designed to bind to two liver-specific factor binding sites in the Enhancer I core region of HBV genomic DNA. Hepatocyte Nuclear Factor 3 (HNF3) and Hepatocyte Nuclear Factor 4 (HNF4) bind to sites in the core region, with the HNF3 site being 5' to the HNF4 site. The HNF3 and HNF4 sites overlap or are adjacent to binding sites for a number of more ubiquitous factors, and are termed nuclear receptor response elements (NRRE). These elements are critical in regulating HBV transcription and replication in infected hepatocytes, with mutations in the HNF3 and HNF4 binding sites having been demonstrated to greatly reduce the levels of HBV replication (Bock *et al.*, 2000, *J. Virology*, 74, 2193)

Oligonucleotides (**Table XV**) were designed to bind to either the positive or negative strands of the HNF3 or HNF4 binding sites. Scrambled controls were made to match each oligo. Each oligo was synthesized in all 2'-O-methyl/all phosphorothioate, or all 2'-O-allyl/all phosphorothioate chemistries. The initial screening of the oligos was done in the HBsAg transfection/ELISA system in Hep G2 cells. RPI.25654, which targets the negative strand of the HNF4 binding site, shows greater activity in reducing HBsAg levels as compared to RPI.25655, which targets the HNF4 site positive strand, and the scrambled control RPI.25656. This result was observed at both 200 and 400 nM (**Figures 18 and 19**).

In a follow-up study, RPI.25654 reduced HBsAg levels in a dose-dependent manner, from 50-200 nM (**Figure 20**).

Example 15: Transfection of HepG2 Cells with psHBV-1 and Nucleic acid

The human hepatocellular carcinoma cell line Hep G2 was grown in Dulbecco's modified Eagle media supplemented with 10% fetal calf serum, 2 mM glutamine, 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, 25 mM Hepes, 100 units penicillin, and 100 μg/ml streptomycin. To generate a replication competent cDNA, prior to transfection the HBV genomic sequences are excised from the bacterial plasmid sequence contained in the psHBV-1 vector This was done with an EcoRI and Hind III restriction digest. Following completion of the digest, a ligation was performed under dilute conditions (20 μg/ml) to favor intermolecular ligation. The total ligation mixture was then concentrated using Qiagen spin columns. One skilled in the art would realize that other methods can be used to generate a replication competent cDNA

Secreted alkaline phosphatase (SEAP) was used to normalize the HBsAg levels to control for transfection variability. The pSEAP2-TK control vector was constructed by ligating a Bgl II-Hind III fragment of the pRL-TK vector (Promega), containing the herpes simplex virus thymidine kinase promoter region, into *Bgl* II/*Hind* III digested pSEAP2-Basic (Clontech). Hep G2 cells were plated (3×10^4 cells/well) in 96-well microtiter plates and incubated overnight. A lipid/DNA/nucleic acid complex was formed containing (at final concentrations) cationic lipid (15 μg/ml), prepared psHBV-1 (4.5 μg/ml), pSEAP2-TK (0.5 μg/ml), and nucleic acid (100 μM). Following a 15 min. incubation at 37° C, the complexes were added to the plated Hep G2 cells. Media was removed from the cells 96 hr. post-transfection for HBsAg and SEAP analysis.

Transfection of the human hepatocellular carcinoma cell line, Hep G2, with replication competent HBV DNA results in the expression of HBV proteins and the production of virions.

Example 16: Analysis of HBsAg and SEAP Levels Following Nucleic Acid Treatment

Immulon 4 (Dynax) microtiter wells were coated overnight at 4° C with anti-HBsAg Mab (Biostride B88-95-31ad,ay) at 1 μg/ml in Carbonate Buffer (Na2CO3 15 mM, NaHCO3 35 mM, pH 9.5). The wells were then washed 4x with PBST (PBS, 0.05% Tween® 20) and blocked for 1 hr at 37° C with PBST, 1% BSA. Following washing as above, the wells were dried at 37° C for 30 min. Biotinylated goat anti-HBsAg (Accurate YVS1807) was diluted 1:1000 in PBST and incubated in the wells for 1 hr. at 37° C. The wells were washed 4x with PBST. Streptavidin/Alkaline Phosphatase Conjugate (Pierce 21324) was diluted to 250

ng/ml in PBST, and incubated in the wells for 1 hr. at 37° C. After washing as above, p-nitrophenyl phosphate substrate (Pierce 37620) was added to the wells, which were then incubated for 1 hr. at 37° C. The optical density at 405 nm was then determined. SEAP levels were assayed using the Great EscAPe® Detection Kit (Clontech K2041-1), as per the manufacturers instructions.

Example 17: Analysis of HBV DNA expression a HepG2.2.15 murine model

The development of new antiviral agents for the treatment of chronic Hepatitis B has been aided by the use of animal models that are permissive to replication of related Hepadnaviridae such as Woodchuck Hepatitis Virus (WHV) and Duck Hepatitis Virus (DHV). In addition, the use of transgenic mice has also been employed. The human hepatoblastoma cell line, HepG2.2.15, implanted as a subcutaneous (SC) tumor, can be used to produce Hepatitis B viremia in mice. This model is useful for evaluating new HBV therapies. Mice bearing HepG2.2.15 SC tumors show HBV viremia. HBV DNA can be detected in serum beginning on Day 35. Maximum serum viral levels reach $1.9x10^5$ copies/mL by day 49. A study also determined that the minimum tumor volume associated with viremia was 300 mm^3. Therefore, the HepG2.2.15 cell line grown as a SC tumor produces a useful model of HBV viremia in mice. This new model can be suitable for evaluating new therapeutic regimens for chronic Hepatitis B.

HepG2.2.15 tumor cells contain a slightly truncated version of viral HBV DNA and sheds HBV particles. The purpose of this study was to identify what time period viral particles are shed from the tumor. Serum was analyzed for presence of HBV DNA over a time course after HepG2.2.15 tumor inoculation in Athymic Ncr nu/nu mice. HepG2.2.15 cells were carried and expanded in DMEM/10% FBS/2.4% HEPES/1% NEAA/1% Glutamine/1% Sodium Pyruvate media. Cells were resuspended in Delbecco's PBS with calcium/magnesium for injection. One hundred microliters of the tumor cell suspension (at a concentration of 1x108 cells/mL) were injected subcutaneously in the flank of NCR nu/nu female mice with a 23gl needle and 1 cc syringe, thereby giving each mouse $1x10^7$ cells. Tumors were allowed to grow for a period of up to 49 days post tumor cell inoculation. Serum was sampled for analysis on days 1, 7, 14, 35, 42 and 49 post tumor inoculation. Length and width measurements from each tumor were obtained three times per week using a Jamison microcaliper. Tumor volumes were calculated from tumor length/width measurements (tumor volume = $0.5[a(b)^2]$ where a = longest axis of the tumor and b = shortest axis of the tumor). Serum was analyzed for the presence of HBV DNA by the Roche Amplicor HBV moniter TM DNA assay.

Experiment 1

HepG2.2.15 cells were carried and expanded in DMEM/10% FBS/2.4%HEPES/1%NEAA/1% Glutamine/1% Sodium Pyruvate media. Cells were resuspended in Delbecco's PBS with calcium/magnesium for injection. One hundred microliters of the tumor cell suspension (at a concentration of 1x108 cells/mL) were injected subcutaneously in the flank of NCR nu/nu female mice with a 23gl needle and 1 cc syringe, thereby giving each mouse $1x10^7$ cells. Tumors were allowed to grow for a period of up to 49 days post tumor cell inoculation. Serum was sampled for analysis on days 1, 7, 14, 35, 42 and 49 post tumor inoculation. Length and width measurements from each tumor were obtained three times per week using a Jamison microcaliper. Tumor volumes were calculated from tumor length/width measurements (tumor volume = $0.5[a(b)^2]$ where a = longest axis of the tumor and b = shortest axis of the tumor). Serum was analyzed for the presence of HBV DNA by the Roche Amplicor HBV moniter TM DNA assay.

Results

When athymic nu/nu female mice are subcutaneously injected with HepG2.2.15 cells and form tumors, HBV DNA is detected in serum (peak serum level was $1.9x10^5$ copies/mL). There is a positive correlation ($rs = 0.7$, $p < 0.01$) between tumor weight (milligrams) and HB viral copies/mL serum. **Figure 21** shows a plot of HepG2.2.15 tumors in nu/nu female mice as tumor volume vs time. **Table XVI** shows the concentration of HBV DNA in relation to tumor size in the HepG2.2.15 implanted nu/nu female mice used in the study.

Experiment 2

HepG2.2.15 cells were carried and expanded in DMEM/10% FBS/2.4%HEPES/1%NEAA/1% Glutamine/1% Sodium Pyruvate media containing 400 µg/ml G418 antibiotic. G418-resistant cells were resuspended in Dulbecco's PBS with calcium/magnesium for injection. One hundred microliters of the tumor cell suspension (at a concentration of 1x108 cells/mL) were injected subcutaneously in the flank of NCR nu/nu female mice with a 23gl needle and 1 cc syringe, thereby giving each mouse $1x10^7$ cells. Tumors were allowed to grow for a period of up to 49 days post tumor cell inoculation. Serum was sampled for analysis on day 37 post tumor inoculation. Length and width measurements from each tumor were obtained three times per week using a Jamison microcaliper. Tumor volumes were calculated from tumor length/width measurements (tumor volume = $0.5[a(b)^2]$ where a = longest axis of the tumor and b = shortest axis of the tumor). Serum was analyzed for the presence of HBV DNA by the Roche Amplicor HBV moniter TM DNA assay.

Results

When athymic nu/nu female mice are subcutaneously injected with G418 antibiotic resistant HepG2.2.15 cells and form tumors, HBV DNA is detected in serum (peak serum level was 4.0×10^5 copies/mL). There is a positive correlation (rs = 0.7, $p < 0.01$) between tumor weight (milligrams) and HB viral copies/mL serum. **Figure 22** shows a plot of HepG2.2.15 tumors in nu/nu female mice as tumor volume vs time. **Table XVII**shows the concentration of HBV DNA in relation to tumor size in the G418 antibiotic resistant HepG2.2.15 implanted nu/nu female mice used in the study.

Example 18: Identification of Potential Enzymatic nucleic acid molecules Cleavage Sites in HCV RNA

The sequence of HCV RNA was screened for accessible sites using a computer folding algorithm. Regions of the mRNA that did not form secondary folding structures and contained potential enzymatic nucleic acid cleavage sites were identified. The sequences of these cleavage sites are shown in **Tables XVIII, XIX, XX and XXIII**.

Example 19: Selection of Enzymatic nucleic acid molecules Cleavage Sites in HCV RNA

Enzymatic nucleic acid target sites were chosen by analyzing sequences of Human HCV (Genbank accession Nos: D11168 , D50483.1, L38318 and S82227) and prioritizing the sites on the basis of folding. Enzymatic nucleic acid molecules are designed that could bind each target and are individually analyzed by computer folding (Christoffersen *et al.,* 1994 *J. Mol. Struc. Theochem,* 311, 273; Jaeger *et al.*, 1989, *Proc. Natl. Acad. Sci. USA*, 86, 7706) to assess whether the enzymatic nucleic acid molecules sequences fold into the appropriate secondary structure. Those enzymatic nucleic acid molecules with unfavorable intramolecular interactions between the binding arms and the catalytic core can be eliminated from consideration. As noted below, varying binding arm lengths can be chosen to optimize activity. Generally, at least 4 bases on each arm are able to bind to, or otherwise interact with, the target RNA.

Example 20: Chemical Synthesis and Purification of Enzymatic nucleic acids

Enzymatic nucleic acid molecules can be designed to anneal to various sites in the RNA message. The binding arms of the enzymatic nucleic acid molecules are complementary to the target site sequences described above. The enzymatic nucleic acid molecules can be chemically synthesized using, for example, RNA syntheses such as those described above and those described in Usman et al., (1987 J. Am. Chem. Soc., 109, 7845), Scaringe et al., (1990 Nucleic Acids Res., 18, 5433) and Wincott et al., supra. Such methods make use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. The average stepwise coupling yields are

typically >98%. Enzymatic nucleic acid molecules can be modified to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-C-allyl, 2'-flouro, 2'-O-methyl, 2'-H (for a review see Usman and Cedergren, 1992 TIBS 17, 34).

Enzymatic nucleic acid molecules can also be synthesized from DNA templates using bacteriophage T7 RNA polymerase (Milligan and Uhlenbeck, 1989, Methods Enzymol. 180, 51). Enzymatic nucleic acid molecules can be purified by gel electrophoresis using known methods, or can be purified by high pressure liquid chromatography (HPLC; See Wincott et al., supra; the totality of which is hereby incorporated herein by reference), and are resuspended in water. The sequences of chemically synthesized enzymatic nucleic acid constructs are shown below in **Tables XX, XXI and XXIII.** The antisense nucleic acid molecules shown in **Table XXII** were chemically synthesized.

Inactive enzymatic nucleic acid molecules, for example inactive hammerhead enzymatic nucleic acids, can be synthesized by substituting the order of G5A6 and substituting a U for A14 (numbering from Hertel et al., 1992 Nucleic Acids Res., 20, 3252).

Example 21: Enzymatic Nucleic Acid Cleavage of HCV RNA Target *in vitro*

Enzymatic nucleic acid molecules targeted to the HCV are designed and synthesized as described above. These enzymatic nucleic acid molecules can be tested for cleavage activity *in vitro*, for example using the following procedure. The target sequences and the nucleotide location within the HCV are given in **Tables XVIII, XIX, XX and XXIII.**

Cleavage Reactions: Full-length or partially full-length, internally-labeled target RNA for enzymatic nucleic acid molecule cleavage assay is prepared by *in vitro* transcription in the presence of [α-^{32}P] CTP, passed over a G 50 Sephadex column by spin chromatography and used as substrate RNA without further purification. Alternately, substrates are 5'-^{32}P-end labeled using T4 polynucleotide kinase enzyme. Assays are performed by pre-warming a 2X concentration of purified enzymatic nucleic acid molecule in enzymatic nucleic acid molecule cleavage buffer (50 mM Tris-HCl, pH 7.5 at 37°C, 10 mM $MgCl_2$) and the cleavage reaction was initiated by adding the 2X enzymatic nucleic acid molecule mix to an equal volume of substrate RNA (maximum of 1-5 nM) that was also pre-warmed in cleavage buffer. As an initial screen, assays are carried out for 1 hour at 37°C using a final concentration of either 40 nM or 1 mM enzymatic nucleic acid molecule, *i.e.*, enzymatic nucleic acid molecule excess. The reaction is quenched by the addition of an equal volume of 95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol after which the sample is heated to 95°C for 2 minutes, quick chilled and loaded onto a denaturing polyacrylamide gel. Substrate RNA and the specific RNA cleavage products generated by enzymatic nucleic acid molecule cleavage are visualized on an autoradiograph of the gel. The

percentage of cleavage is determined by Phosphor Imager® quantitation of bands representing the intact substrate and the cleavage products.

Alternatively, enzymatic nucleic acid molecules and substrates were synthesized in 96-well format using 0.2μmol scale. Substrates were 5'-^{32}P labeled and gel purified using 7.5% polyacrylamide gels, and eluting into water. Assays were done by combining trace substrate with 500nM enzymatic nucleic acid or greater, and initiated by adding final concentrations of 40mM Mg^{+2}, and 50mM Tris-Cl pH 8.0. For each enzymatic nucleic acid/substrate combination a control reaction was done to ensure cleavage was not the result of non-specific substrate degradation. A single three hour time point was taken and run on a 15% polyacrylamide gel to asses cleavage activity. Gels were dried and scanned using a Molecular Dynamics Phosphorimager and quantified using Molecular Dynamics ImageQuant software. Percent cleaved was determined by dividing values for cleaved substrate bands by full-length (uncleaved) values plus cleaved values and multiplying by 100 (%cleaved=[C/(U+C)]*100). In vitro cleavage data of enzymatic nucleic acid molecules targeting plus and minus strand HCV RNA is shown in **Table XXIII**.

Example 22: Inhibition of Luciferase Activity Using HCV Targeting Enzymatic nucleic acids in OST7 Cells

The capability of enzymatic nucleic acids to inhibit HCV RNA intracellularly was tested using a dual reporter system that utilizes both firefly and Renilla luciferase (**Figure 23**). The enzymatic nucleic acids targeted to the 5' HCV UTR region, which when cleaved, would prevent the translation of the transcript into luciferase.

Synthesis of Stabilized Enzymatic nucleic acids

Enzymatic nucleic acids were designed to target 15 sites within the 5'UTR of the HCV RNA (**Figure 24**) and synthesized as previously described, except that all enzymatic nucleic acids contain two 2'-amino uridines. Enzymatic nucleic acid and paired control sequences for targeted sites used in various examples herein are shown in **Table XXI**.

Reporter plasmids

The T7/HCV/firefly luciferase plasmid (HCVT7C$_{1\text{-}341}$, genotype 1a) was graciously provided by Aleem Siddiqui (University of Colorado Health Sciences Center, Denver, CO). The T7/HCV/firefly luciferase plasmid contains a T7 bacteriophage promoter upstream of the HCV 5'UTR (nucleotides 1-341)/firefly luciferase fusion DNA. The Renilla luciferase control plasmid (pRLSV40) was purchased from PROMEGA.

Luciferase assay

Dual luciferase assays were carried out according to the manufacturer's instructions (PROMEGA) at 4 hours after co-transfection of reporter plasmids and enzymatic nucleic acids. All data is shown as the average ratio of HCV/firefly luciferase luminescence over Renilla luciferase luminescence as determined by triplicate samples $\pm$ SD.

Cell culture and transfections

OST7 cells were maintained in Dulbecco's modified Eagle's medium (GIBCO BRL) supplemented with 10% fetal calf serum, L-glutamine (2 mM) and penicillin/streptomycin. For transfections, OST7 cells were seeded in black-walled 96-well plates (Packard) at a density of 12,500 cells/well and incubated at 37°Cunder 5% CO_2 for 24 hours. Co-transfection of target reporter HCVT7C (0.8 μg/mL), control reporter pRLSV40, (1.2 μg/mL) and enzymatic nucleic acid, (50 - 200 nM) was achieved by the following method: a 5X mixture of HCVT7C (4 μg/mL), pRLSV40 (6 μg/mL) enzymatic nucleic acid (250 – 1000 nM) and cationic lipid (28.5 μg/mL) was made in 150 μL of OPTI-MEM (GIBCO BRL) minus serum. Reporter/enzymatic nucleic acid/lipid complexes were allowed to form for 20 min at 37°Cunder 5% CO_2. Medium was aspirated from OST7 cells and replaced with 120 μL of OPTI-MEM (GIBCO BRL) minus serum, immediately followed by the addition of 30 μL of 5X reporter/enzymatic nucleic acid/lipid complexes. Cells were incubated with complexes for 4 hours at 37°Cunder 5% CO_2.

IC50 determinations for dose response curves

Apparent IC_{50} values were calculated by linear interpolation. The apparent IC_{50} is 1/2 the maximal response between the two consecutive points in which approximately 50% inhibition of HCV/luciferase expression is observed on the dose curve.

Quantitation of RNA Samples

Total RNA from transfected cells was purified using the Qiagen RNeasy 96 procedure including a DNase I treatment according to the manufacturer's instructions. Real time RT-PCR (Taqman assay) was performed on purified RNA samples using separate primer/probe sets specific for either firefly or Renilla luciferase RNA. Firefly luciferase primers and probe were upper (5'-CGGTCGGTAAAGTTGTTCCATT-3') (SEQ ID NO. **16202**), lower (5'-CCTCTGACACATAATTCGCCTCT-3') (SEQ ID NO. **16203**), and probe (5'-FAM-TGAAGCGAAGGTTGTGGATCTGGATACC-TAMRA-3') (SEQ ID NO **16204**), and Renilla luciferase primers and probe were upper (5'-GTTTATTGAATCGGACCCAGGAT-3') (SEQ ID NO. **16205**), lower (5'-AGGTGCATCTTCTTGCGAAAA-3') (SEQ ID NO. **16206**), and probe (5'-FAM-CTTTTCCAATGCTATTGTTGAAGGTGCCAA-3') (SEQ ID NO. **16207**) -TAMRA, both sets of primers and probes were purchased from Integrated DNA

Technologies. RNA levels were determined from a standard curve of amplified RNA purified from a large-scale transfection. RT minus controls established that RNA signals were generated from RNA and not residual plasmid DNA. RT-PCR conditions were: 30 min at 48°C, 10 min at 95°C, followed by 40 cycles of 15 sec at 95°C and 1 min at 60°C. Reactions were performed on an ABI Prism 7700 sequence detector. Levels of firefly luciferase RNA were normalized to the level of Renilla luciferase RNA present in the same sample. Results are shown as the average of triplicate treatments ± SD.

Example 23: Inhibition of HCV 5'UTR-luciferase expression by synthetic stabilized enzymatic nucleic acids

The primary sequence of the HCV 5'UTR and characteristic secondary structure (**Figure 24**) is highly conserved across all HCV genotypes, thus making it a very attractive target for enzymatic nucleic acid-mediated cleavage. Enzymatic hammerhead nucleic acids, as a generally shown in **Figure 25** and **Table XXI** (RPI 12249-12254, 12257-12265) were designed and synthesized to target 15 of the most highly conserved sites in the 5'UTR of HCV RNA. These synthetic enzymatic nucleic acids were stabilized against nuclease degradation by the addition of modifications such as 2'-*O*-methyl nucleotides, 2'-amino-uridines at U4 and U7 core positions, phosphorothioate linkages, and a 3'-inverted abasic cap.

In order to mimic cytoplasmic transcription of the HCV genome, OST7 cells were transfected with a target reporter plasmid containing a T7 bacteriophage promoter upstream of a HCV 5'UTR/firefly luciferase fusion gene. Cytoplasmic expression of the target reporter is facilitated by high levels of T7 polymerase expressed in the cytoplasm of OST7 cells. Co-transfection of target reporter HCVT7C$_{1\text{-}341}$ (firefly luciferase), control reporter pRLSV40 (Renilla luciferase) and enzymatic nucleic acid was carried out in the presence of cationic lipid. To determine the background level of luciferase activity, applicant used a control enzymatic nucleic acid that targets an irrelevant, non-HCV sequence. Transfection of reporter plasmids in the presence of this irrelevant control enzymatic nucleic acid (ICR) resulted in a slight decrease of reporter expression when compared to transfection of reporter plasmids alone. Therefore, the ICR was used to control for non-specific effects on reporter expression during treatment with HCV specific enzymatic nucleic acids. Renilla luciferase expression from the pRLSV40 reporter was used to normalize for transfection efficiency and sample recovery.

Of the 15 amino-modified hammerhead enzymatic nucleic acids tested, 12 significantly inhibited HCV/luciferase expression (> 45%, $P < 0.05$) as compared to the ICR (**Figure 26A**). These data suggest that most of the HCV 5'UTR sites targeted here are accessible to enzymatic nucleic acid binding and subsequent RNA cleavage. To investigate further the

enzymatic nucleic acid-dependent inhibition of HCV/luciferase activity, hammerhead enzymatic nucleic acids designed to cleave after sites 79, 81, 142, 192, 195, 282 or 330 of the HCV 5'UTR were selected for continued study because their anti-HCV activity was the most efficacious over several experiments. A corresponding attenuated core (AC) control was synthesized for each of the 7 active enzymatic nucleic acids (**Table XX**). Each paired AC control contains similar nucleotide composition to that of its corresponding active enzymatic nucleic acid however, due to scrambled binding arms and changes to the catalytic core, lacks the ability to bind or catalyze the cleavage of HCV RNA. Treatment of OST7 cells with enzymatic nucleic acids designed to cleave after sites 79, 81, 142, 195 or 330 resulted in significant inhibition of HCV/luciferase expression (65%, 50%, 50%, 80% and 80%, respectively) when compared to HCV/luciferase expression in cells treated with corresponding ACs, $P < 0.05$ (**Figure 26B**). It should be noted that treatment with either the ICR or ACs for sites 79, 81, 142 or 192 caused a greater reduction of HCV/luciferase expression than treatment with ACs for sites 195, 282 or 330. The observed differences in HCV/luciferase expression after treatment with ACs most likely represents the range of activity due to non-specific effects of oligonucleotide treatment and/or differences in base composition. Regardless of differences in HCV/luciferase expression levels observed as a result of treatment with ACs, active enzymatic nucleic acids designed to cleave after sites 79, 81, 142, 195, or 330 demonstrated similar and potent anti-HCV activity (**Figure 26B**).

Example 24: Synthetic stabilized enzymatic nucleic acids inhibit HCV/luciferase expression in a concentration-dependent manner

In order to characterize enzymatic nucleic acid efficacy in greater detail, these same 5 lead hammerhead enzymatic nucleic acids were tested for their ability to inhibit HCV/luciferase expression over a range of enzymatic nucleic acid concentrations (0 nM - 100 nM). For constant transfection conditions, the total concentration of nucleic acid was maintained at 100 nM for all samples by mixing the active enzymatic nucleic acid with its corresponding AC. Moreover, mixing of active enzymatic nucleic acid and AC maintains the lipid to nucleic acid charge ratio. A concentration-dependent inhibition of HCV/luciferase expression was observed after treatment with each of the 5 enzymatic nucleic acids (**Figures 27A-E**). By linear interpolation, the enzymatic nucleic acid concentration resulting in 50% inhibition (apparent IC_{50}) of HCV/luciferase expression ranged from 40 - 215 nM. The two most efficacious enzymatic nucleic acids were those designed to cleave after sites 195 or 330 with apparent IC_{50} values of 46 nM and 40 nM, respectively (**Figures 27D and E**).

Example 25: An enzymatic nucleic acid mechanism is required for the observed inhibition of HCV/luciferase expression

To confirm that an enzymatic nucleic acid mechanism of action was responsible for the observed inhibition of HCV/luciferase expression, paired binding-arm attenuated core (BAC) controls (RPI 15291 and 15294) were synthesized for direct comparison to enzymatic nucleic acids targeting sites 195 (RPI 12252) and 330 (RPI 12254). Paired BACs can specifically bind HCV RNA but are unable to promote RNA cleavage because of changes in the catalytic core and, thus, can be used to assess inhibition due to binding alone. Also included in this comparison were paired SAC controls (RPI 15292 and 15295) that contain scrambled binding arms and attenuated catalytic cores, and so lack the ability to bind the target RNA or to catalyze target RNA cleavage.

Enzymatic nucleic acid cleavage of target RNA should result in both a lower level of HCV/luciferase RNA and a subsequent decrease in HCV/luciferase expression. In order to analyze target RNA levels, a reverse transcriptase/polymerase chain reaction (RT-PCR) assay was employed to quantify HCV/luciferase RNA levels. Primers were designed to amplify the luciferase coding region of the HCV 5'UTR/luciferase RNA. This region was chosen because HCV-targeted enzymatic nucleic acids that might co-purify with cellular RNA would not interfere with RT-PCR amplification of the luciferase RNA region. Primers were also designed to amplify the Renilla luciferase RNA so that Renilla RNA levels could be used to control for transfection efficiency and sample recovery.

OST7 cells were treated with active enzymatic nucleic acids designed to cleave after sites 195 or 330, paired SACs, or paired BACs. Treatment with enzymatic nucleic acids targeting site 195 or 330 resulted in a significant reduction of HCV/luciferase RNA when compared to their paired SAC controls ($P < 0.01$). In this experiment the site 195 enzymatic nucleic acid was more efficacious than the site 330 enzymatic nucleic acid (**Figure 28A**). Treatment with paired BACs that target site 195 or 330 did not reduce HCV/luciferase RNA when compared to the corresponding SACs, thus confirming that the ability to bind alone does not result in a reduction of HCV/luciferase RNA.

To confirm that enzymatic nucleic acid-mediated cleavage of target RNA is necessary for inhibition of HCV/luciferase expression, HCV/luciferase activity was determined in the same experiment. As expected, significant inhibition of HCV/luciferase expression was observed after treatment with active enzymatic nucleic acids when compared to paired SACs (**Figure 28B**). Importantly, treatment with paired BACs did not inhibit HCV/luciferase expression, thus confirming that the ability to bind alone is also not sufficient to inhibit translation. As observed in the RNA assay, the site 195 enzymatic nucleic acid was more efficacious than the site 330 enzymatic nucleic acid in this experiment. However, a correlation between enzymatic nucleic acid-mediated HCV RNA reduction and inhibition of HCV/luciferase translation was observed for enzymatic nucleic acids to both sites. The

reduction in target RNA and the necessity for an active enzymatic nucleic acid catalytic core confirm that a enzymatic nucleic acid mechanism is required for the observed reduction in HCV/luciferase protein activity in cells treated with site 195 or site 330 enzymatic nucleic acids.

Example 26: Zinzyme Inhibition of chimeric HCV/Poliovirus replication

During HCV infection, viral RNA is present as a potential target for enzymatic nucleic acid cleavage at several processes: un-coating, translation, RNA replication and packaging. Target RNA can be more or less accessible to enzymatic nucleic acid cleavage at any one of these steps. Although the association between the HCV initial ribosome entry site (IRES) and the translation apparatus is mimicked in the HCV 5'UTR/luciferase reporter system, these other viral processes are not represented in the OST7 system. The resulting RNA/protein complexes associated with the target viral RNA are also absent. Moreover, these processes can be coupled in an HCV-infected cell which could further impact target RNA accessibility. Therefore, applicant tested whether enzymatic nucleic acids designed to cleave the HCV 5'UTR could effect a replicating viral system.

Recently, Lu and Wimmer characterized a HCV-poliovirus chimera in which the poliovirus IRES was replaced by the IRES from HCV (Lu & Wimmer, 1996, Proc. Natl. Acad. Sci. USA. 93, 1412-1417). Poliovirus (PV) is a positive strand RNA virus like HCV, but unlike HCV is non-enveloped and replicates efficiently in cell culture. The HCV-PV chimera expresses a stable, small plaque phenotype relative to wild type PV.

The following enzymatic nucleic acid molecules (zinzymes) were synthesized and tested for replicative inhibition of an HCV/Poliovirus chimera: RPI 18763, RPI 18812, RPI 18749, RPI 18765, RPI 18792, and RPI 18814 **(Table XX)**. A scrambled attenuated core enzymatic nucleic acid, RPI 18743, was used as a control.

HeLa cells were infected with the HCV-PV chimera for 30 minutes and immediately treated with enzymatic nucleic acid. HeLa cells were seeded in U-bottom 96-well plates at a density of 9000-10,000 cells/well and incubated at 37°C under 5% CO2 for 24 h. Transfection of nucleic acid (200 nM) was achieved by mixing of 10X nucleic acid (2000 nM) and 10X of a cationic lipid (80 μg/ml) in DMEM (Gibco BRL) with 5% fetal bovine serum (FBS). Nucleic acid/lipid complexes were allowed to incubate for 15 minutes at 37°C under 5% CO2. Medium was aspirated from cells and replaced with 80 μl of DMEM (Gibco BRL) with 5% FBS serum, followed by the addition of 20 μls of 10X complexes. Cells were incubated with complexes for 24 hours at 37°C under 5% CO2 .

The yield of HCV-PV from treated cells was quantified by plaque assay. The plaque assays were performed by diluting virus samples in serum-free DMEM (Gibco BRL) and applying 100 μl to HeLa cell monolayers (~80% confluent) in 6-well plates for 30 minutes. Infected monolayers were overlayed with 3 ml 1.2% agar (Sigma) and incubated at 37°C under 5% CO2. Two or three days later the overlay was removed, monolayers were stained with 1.2% crystal violet, and plaque forming units were counted. The results for the zinzyme inhibition of HCV-PV replication are shown in **Figure 33**.

Example 27: Antisense inhibition of chimeric HCV/Poliovirus replication

Antisense nucleic acid molecules (RPI 17501 and RPI 17498, **Table XXII**) were tested for replicative inhibition of an HCV/Poliovirus chimera compared to scrambled controls. An antisense nucleic acid molecule is a non-enzymatic nucleic acid molecule that binds to target RNA by means of RNA-RNA or RNA-DNA or RNA-PNA (protein nucleic acid; Egholm et al., 1993 Nature 365, 566) interactions and alters the activity of the target RNA (for a review, see Stein and Cheng, 1993 Science 261, 1004 and Woolf et al., US patent No. 5,849,902). Typically, antisense molecules are complementary to a target sequence along a single contiguous sequence of the antisense molecule. However, in certain embodiments, an antisense molecule can bind to substrate such that the substrate molecule forms a loop, and/or an antisense molecule can bind such that the antisense molecule forms a loop. Thus, the antisense molecule can be complementary to two (or even more) non-contiguous substrate sequences or two (or even more) non-contiguous sequence portions of an antisense molecule can be complementary to a target sequence or both. For a review of current antisense strategies, see Schmajuk et al., 1999, J. Biol. Chem., 274, 21783-21789, Delihas et al., 1997, Nature, 15, 751-753, Stein et al., 1997, Antisense N. A. Drug Dev., 7, 151, Crooke, 2000, Methods Enzymol., 313, 3-45; Crooke, 1998, Biotech. Genet. Eng. Rev., 15, 121-157, Crooke, 1997, Ad. Pharmacol., 40, 1-49. In addition, antisense DNA can be used to target RNA by means of DNA-RNA interactions, thereby activating RNase H, which digests the target RNA in the duplex. The antisense oligonucleotides can comprise one or more RNAse H activating region, which is capable of activating RNAse H cleavage of a target RNA. Antisense DNA can be synthesized chemically or expressed via the use of a single stranded DNA expression vector or equivalent thereof. Additionally, antisense molecules can be used in combination with the enzymatic nucleic acid molecules of the instant invention.

A RNase H activating region is a region (generally greater than or equal to 4-25 nucleotides in length, preferably from 5-11 nucleotides in length) of a nucleic acid molecule capable of binding to a target RNA to form a non-covalent complex that is recognized by cellular RNase H enzyme (see for example Arrow et al., US 5,849,902; Arrow et al., US 5,989,912). The RNase H enzyme binds to the nucleic acid molecule-target RNA complex

and cleaves the target RNA sequence. The RNase H activating region comprises, for example, phosphodiester, phosphorothioate (preferably at least four of the nucleotides are phosphorothiote substitutions; more specifically, 4-11 of the nucleotides are phosphorothiote substitutions); phosphorodithioate, 5'-thiophosphate, or methylphosphonate backbone chemistry or a combination thereof. In addition to one or more backbone chemistries described above, the RNase H activating region can also comprise a variety of sugar chemistries. For example, the RNase H activating region can comprise deoxyribose, arabino, fluoroarabino or a combination thereof, nucleotide sugar chemistry. Those skilled in the art will recognize that the foregoing are non-limiting examples and that any combination of phosphate, sugar and base chemistry of a nucleic acid that supports the activity of RNase H enzyme is within the scope of the definition of the RNase H activating region and the instant invention.

HeLa cells were infected with the HCV-PV chimera for 30 minutes and immediately treated with antisense nucleic acid. HeLa cells were seeded in U-bottom 96-well plates at a density of 9000-10,000 cells/well and incubated at 37°C under 5% CO2 for 24 h. Transfection of nucleic acid (200 nM) was achieved by mixing of 10X nucleic acid (2000 nM) and 10X of a cationic lipid (80 μg/ml) in DMEM (Gibco BRL) with 5% fetal bovine serum (FBS). Nucleic acid/lipid complexes were allowed to incubate for 15 minutes at 37°C under 5% CO2. Medium was aspirated from cells and replaced with 80 μl of DMEM (Gibco BRL) with 5% FBS serum, followed by the addition of 20 μls of 10X complexes. Cells were incubated with complexes for 24 hours at 37°C under 5% CO2 .

The yield of HCV-PV from treated cells was quantified by plaque assay. The plaque assays were performed by diluting virus samples in serum-free DMEM (Gibco BRL) and applying 100 μl to HeLa cell monolayers (~80% confluent) in 6-well plates for 30 minutes. Infected monolayers were overlayed with 3 ml 1.2% agar (Sigma) and incubated at 37°C under 5% CO2. Two or three days later the overlay was removed, monolayers were stained with 1.2% crystal violet, and plaque forming units were counted. The results for the antisense inhibition of HCV-PV are shown in **Figure 34**.

Example 28: Nucleic acid Inhibition of Chimeric HCV/PV in combination with Interferon

One of the limiting factors in interferon (IFN) therapy for chronic HCV are the toxic side effects associated with IFN. Applicant has reasoned that lowering the dose of IFN needed can reduce these side effects. Applicant has previously shown that enzymatic nucleic acid molecules targeting HCV RNA have a potent antiviral effect against replication of an HCV-poliovirus (PV) chimera (Macejak *et al.*, 2000, *Hepatology*, 31, 769-776). In order to determine if the antiviral effect of type 1 IFN could be improved by the addition of anti-HCV enzymatic nucleic acid treatment, a dose response (0 U/ml to 100 U/ml) with IFN alfa 2a or

IFN alfa 2b was performed in HeLa cells in combination with 200 nM site 195 anti-HCV enzymatic nucleic acid (RPI 13919) or enzymatic nucleic acid control (SAC) treatment. The SAC control (RPI 17894) is a scrambled binding arm, attenuated core version of the site 195 enzymatic nucleic acid (RPI 13919). IFN dose responses were performed with different pretreatment regimes to find the dynamic range of inhibition in this system. In these studies, HeLa cells were used instead of HepG2 because of more efficient enzymatic nucleic acid delivery (Macejak *et al.*, 2000, *Hepatology,* 31, 769-776).

Cells and Virus

HeLa cells were maintained in DMEM (BioWhittaker, Walkersville, MD) supplemented with 5% fetal bovine serum. A cloned DNA copy of the HCV-PV chimeric virus was a gift of Dr. Eckard Wimmer (NYU, Stony Brook, NY). An RNA version was generated by in vitro transcription and transfected into HeLa cells to produce infectious virus (Lu and Wimmer, 1996, PNAS USA., 93, 1412-1417).

Enzymatic nucleic acid Synthesis

Nuclease resistant enzymatic nucleic acids and control oligonucleotides containing 2'-O-methyl-nucleotides, 2'-deoxy-2'-C-allyl uridine, a 3'-inverted abasic cap, and phosphorothioate linkages were chemically synthesized. The anti-HCV enzymatic nucleic acid (RPI 13919) targeting cleavage after nucleotide 195 of the 5' UTR of HCV is shown in **Table XX**. Attenuated core controls have nucleotide changes in the core sequence that greatly diminished the enzymatic nucleic acid's cleavage activity. The attenuated controls either contain scrambled binding arms (referred to as SAC, RPI 18743) or maintain binding arms (BAC, RPI 17894) capable of binding to the HCV RNA target.

Enzymatic nucleic acid Delivery

A cationic lipid was used as a cytofectin agent. HeLa cells were seeded in 96-well plates at a density of 9000-10,000 cells/well and incubated at 37°Cunder 5% CO2 for 24 h. Transfection of enzymatic nucleic acid or control oligonucleotides (200 nM) was achieved by mixing 10X enzymatic nucleic acid or control oligonucleotides (2000 nM) with 10X RPI.9778 (80 μg/ml) in DMEM containing 5% fetal bovine serum (FBS) in U-bottom 96-well plates to make 5X complexes. Enzymatic nucleic acid/lipid complexes were allowed to incubate for 15 min at 37°C under 5% CO2. Medium was aspirated from cells and replaced with 80 μl of DMEM (Gibco BRL) containing 5% FBS serum, followed by the addition of 20 μl of 5X complexes. Cells were incubated with complexes for 24 h at 37°Cunder 5% CO2.

Interferon/Enzymatic nucleic acid Combination Treatment

Interferon alfa 2a (Roferon®) was purchased from Roche Bioscience (Palo Alto, CA). Interferon alfa 2b (Intron A®) was purchased from Schering-Plough Corporation (Madison, NJ). Consensus interferon (interferon-alfa-con 1) was a generous gift of Amgen, Inc. (Thousand Oaks, CA). For the basis of comparison, the manufacturers' specified units were used in the studies reported here; however, the manufacturers' unit definitions of these three IFN preparations are not necessarily the same. Nevertheless, since clinical dosing is based on the manufacturers' specified units, a direct comparison based on these units has relevance to clinical therapeutic indices. HeLa cells were seeded (10,000 cells per well) and incubated at 37°Cunder 5% CO2 for 24 h. Cells were then pre-treated with interferon in complete media (DMEM + 5% FBS) for 4 h and then infected with HCV-PV at a multiplicity of infection (MOI) = 0.1 for 30 min. The viral inoculum was then removed and enzymatic nucleic acid or attenuated control (SAC or BAC) was delivered with the cytofectin formulation (8 μg/ml) in complete media for 24 h as described above. Where indicated for enzymatic nucleic acid dose response studies, active enzymatic nucleic acid was mixed with SAC to maintain a 200 nM total oligonucleotide concentration and the same lipid charge ratio. After 24 h, cells were lysed to release virus by three cycles of freeze/thaw. Virus was quantified by plaque assay and viral yield is reported as mean plaque forming units per ml (pfu/ml) + SD. All experiments were repeated at least twice and the trends in the results reported were reproducible. Significance levels (P values) were determined by the Student's test.

Plaque Assay

Virus samples were diluted in serum-free DMEM and 100 μl applied to Vero cell monolayers (~80% confluent) in 6-well plates for 30 min. Infected monolayers were overlaid with 3 ml 1.2% agar (Sigma Chemical Company, St. Louis, MO) and incubated at 37°Cunder 5% CO2. When plaques were visible (after two to three days) the overlay was removed, monolayers were stained with 1.2% crystal violet, and plaque forming units were counted.

Results

As shown in **Figure 29A** and **29B**, treatment with the site 195 (RPI 13919) anti-HCV hammerhead enzymatic nucleic acid alone (0 U/ml IFN) resulted in viral replication that was dramatically reduced compared to SAC-treated cells (85%, P<0.01). For both IFN alfa 2a (**Figure 29A**) or IFN alfa 2b (**Figure 29B**), treatment with 25 U/ml resulted in a ~90% inhibition of HCV-PV replication in SAC-treated cells as compared to cells treated with SAC alone (p<0.01 for both observations). The maximal level of inhibition in SAC-treated cells (94%) was achieved by treatment with ≥50U/ml of either IFN alfa 2a or IFN alfa 2b (p<0.01 for both observations *versus* SAC alone). Maximal inhibition could however, be achieved by a 5-fold lower dose of IFN alfa 2a (10 U/ml) if enzymatic nucleic acid targeting site 195 in the 5' UTR of HCV RNA was given in combination (**Figure 29A**, p<0.01). While the

additional effect of enzymatic nucleic acid treatment on IFN alfa 2b-treated cells at 10 U/ml was very slight, the combined effect with 25 U/ml IFN alfa 2b was greater in magnitude (**Figure 29B**). For both interferons tested, pretreatment with 25 U/ml in combination with 200 nM site 195 anti-HCV enzymatic nucleic acid resulted in an even greater level of inhibition of viral replication (>98%) compared to replication in cells treated with 200 nM SAC alone ($P<0.01$).

A dose response of the site 195 anti-HCV enzymatic nucleic acid was also performed in HeLa cells, either with or without 12.5 U/ml IFN alfa 2a or IFN alfa 2b pretreatment. As shown in **Figure 30**, enzymatic nucleic acid-mediated inhibition was dose-dependent and a significant inhibition of HCV-PV replication (>75% *versus* 0 nM enzymatic nucleic acid, $P<0.01$) could be achieved by treatment with ≥150 nM anti-HCV enzymatic nucleic acid alone (no IFN). However, in IFN-pretreated cells, the dose of anti-HCV enzymatic nucleic acid needed to achieve this level of inhibition was decreased 3-fold to 50 nM ($P<0.01$ *versus* 0 nM enzymatic nucleic acid). In comparison, treatment with the site 195 anti-HCV enzymatic nucleic acid alone at 50 nM resulted in only ~40% inhibition of virus replication. Pretreatment with IFN enhanced the antiviral effect of site 195 enzymatic nucleic acid at all enzymatic nucleic acid doses, compared to no IFN pretreatment.

Interferon-alfacon1, consensus IFN (CIFN), is another type 1 IFN that is used to treat chronic HCV. To determine if a similar enhancement can occur in CIFN-treated cells, a dose response with CIFN was performed in HeLa cells using 0 U/ml to 12.5 U/ml CIFN in combination with 200 nM site 195 anti-HCV enzymatic nucleic acid or SAC treatment (**Figure 31A**). Again, in the presence of the site 195 anti-HCV enzymatic nucleic acid alone, viral replication was dramatically reduced compared to SAC-treated cells. As shown in **Figure 31A**, treatment with 200 nM anti-HCV enzymatic nucleic acid alone significantly inhibited HCV-PV replication (90% *versus* SAC treatment, $P<0.01$). However, pretreatment with concentrations of CIFN from 1 U/ml to 12.5 U/ml in combination with 200 nM anti-HCV enzymatic nucleic acid resulted in even greater inhibition of viral replication (>98%) compared to replication in cells treated with 200 nM SAC alone ($P<0.01$). It is important to note that pretreatment with 1 U/ml CIFN in SAC-treated cells did not have a significant effect on HCV-poliovirus replication, but in the presence of enzymatic nucleic acid a significant inhibition of replication was observed (>98%, $P<0.01$). Thus, the dose of CIFN needed to achieve a >98% inhibition could be lowered to 1 U/ml in cells also treated with 200 nM site 195 anti-HCV enzymatic nucleic acid.

A dose response of site 195 anti-HCV enzymatic nucleic acid was then performed in HeLa cells, either with or without 12.5 U/ml CIFN pretreatment. As shown in **Figure 31B**, a significant inhibition of HCV-PV replication (>95% *versus* 0 nM enzymatic nucleic acid,

P<0.01) could be achieved by treatment with ≥150 nM anti-HCV enzymatic nucleic acid alone. However, in CIFN-pretreated cells, the dose of anti-HCV enzymatic nucleic acid needed to achieve this level of inhibition was only 50 nM (P<0.01). In comparison, treatment with the site 195 anti-HCV enzymatic nucleic acid alone at 50 nM resulted in ~50% inhibition of virus replication. Thus, as was seen with IFN alfa 2a and IFN alfa 2b, the dose of enzymatic nucleic acid could be reduced 3-fold in the presence of CIFN pretreatment to achieve a similar antiviral effect as enzymatic nucleic acid-treatment alone.

To further explore the combination of lower enzymatic nucleic acid concentration and CIFN, a dose response with 0 U/ml to 12.5 U/ml CIFN was subsequently performed in HeLa cells in combination with 50 nM site 195 anti-HCV enzymatic nucleic acid treatment. In multiple experiments, treatment with 50 nM anti-HCV enzymatic nucleic acid alone inhibited HCV-PV replication 50% – 81% compared to viral replication in SAC-treated cells. As for the experiment shown in **Figure 31A,** treatment with CIFN alone at 5 U/ml resulted in ~50% inhibition of viral replication. However, a four hour pretreatment with 5 U/ml CIFN followed by 50 nM anti-HCV enzymatic nucleic acid treatment resulted in 95% - 97% inhibition compared to SAC-treated cells (P<0.01).

To demonstrate that the enhanced antiviral effect of CIFN and enzymatic nucleic acid combination treatment was dependent upon enzymatic nucleic acid cleavage activity, the effect of CIFN in combination with site 195 anti-HCV enzymatic nucleic acid versus the effect of CIFN in combination with a binding competent, attenuated core, control (BAC) was then compared. The BAC can still bind to its specific RNA target, but is greatly diminished in cleavage activity. Pretreatment with 12.5 U/ml CIFN reduced the viral yield ~90% (7-fold) in cells treated with BAC (compare CIFN versus BAC in **Figure 32**). Cells treated with 200 nM site 195 anti-HCV enzymatic nucleic acid alone produced ~95% (17-fold) less virus than BAC-treated cells (195 RZ BAC in **Figure 32**). The combination of CIFN pretreatment and 200 nM site 195 anti-HCV enzymatic nucleic acid results in an augmented >98% (300-fold) reduction in viral yield (CIFN+RZ versus control in **Figure 32**).

<u>2'-5'-Oligoadenylate Inhibition of HCV</u>

Type 1 Interferon is a key constituent of many effective treatment programs for chronic HCV infection. Treatment with type 1 interferon induces a number of genes and results in an antiviral state within the cell. One of the genes induced is 2', 5' oligoadenylate synthetase, an enzyme that synthesizes short 2', 5' oligoadenylate (2-5A) molecules. Nascent 2-5A subsequently activates a latent RNase, RNase L, which in turn nonspecifically degrades viral RNA. As described herein, ribozymes targeting HCV RNA that inhibit the replication of an HCV-poliovirus (HCV-PV) chimera in cell culture and have shown that this antiviral effect is

augmented if ribozyme is given in combination with type 1 interferon. In addtion, the 2-5A component of the interferon response can also inhibit replication of the HCV-PV chimera.

The antiviral effect of anti-HCV ribozyme treatment is enhanced if type 1 interferon is given in combination. Interferon induces a number of gene products including 2',5' oligoadenylate (2-5A) synthetase, double-stranded RNA-activated protein kinase (PKR), and the Mx proteins. Mx proteins appear to interfere with nuclear transport of viral complexes and are not thought to play an inhibitory role in HCV infection. On the other hand, the additional 2-5A-mediated RNA degradation (via RNase L) and/or the inhibition of viral translation by PKR in interferon-treated cells can augment the ribozyme-mediated inhibition of HCV-PV replication.

To investigate the potential role of the 2-5A/RNase L pathway in this enhancement phenomenon, HCV-PV replication was analyzed in HeLa cells treated exogenously with chemically-synthesized analogs of 2-5A (**Figure 35**), alone and in combination with the anti-HCV ribozyme (RPI 13919). These results were compared to replication in cells treated with interferon and/or anti-HCV ribozyme. Anti-HCV ribozyme was transfected into cells with a cationic lipid. To control for nonspecific effects due to lipid-mediated transfection, a scrambled arm, attenuated core, oligonucleotide (SAC) (RPI 17894) was transfected for comparison. The SAC is the same base composition as the ribozyme but is greatly attenuated in catalytic activity due to changes in the core sequence and cannot bind specifically to the HCV sequence.

As shown in **Figure 36A**, HeLa cells pretreated with 10 U/ml consensus interferon for 4 hours prior to HCV-PV infection resulted in ~70% reduction of viral replication in SAC-treated cells. Similarly, HeLa cells treated with 100 nM anti-HCV ribozyme for 20 hours after infection resulted in an ~80% reduction in viral yield. This antiviral effect was enhanced to ~98% inhibition in HeLa cells pretreated with interferon for 4 hours before infection and then treated with anti-HCV ribozyme for 20 hours after infection. In parallel, a 2-5A compound (analog I, **Figure 35**) that was protected from nuclease digestion at the 3'-end with an inverted abasic moiety was tested. As shown in **Figure 36B**, treatment with 200 nM 2-5A analog I for 4 hours prior to HCV-PV infection only slightly inhibited HCV-PV replication (~20%) in SAC-treated cells. Moreover, the inhibition due to a 20 hour anti-HCV ribozyme treatment was not augmented with a 4 hour pretreatment of 2-5A in combination (compare third bar to fourth bar in **Figure 36B**).

There are several possible possible explanations why the chemically synthesized 2-5A analog was not able to completely activate RNase L. It is possible that the 2-5A analog was not sufficiently stable or that in this experiment the 4 hour pretreatment period was too short for RNase L activation. To test these possibilities, a 2-5A compound containing a 5'-terminal

thiophosphate (P=S) for added nuclease resistance, in addition to the 3'- abasic, was also included (analog II, **Figure 35**). In addition, a longer 2-5A treatment was used. In this experiment (**Figure 37**), HeLa cells were treated with 2-5A or 2-5A(P=S) for 20 hours after HCV-PV infection. Again, anti-HCV ribozyme treatment resulted in >80% inhibition. In contrast to the 20% inhibition of viral replication seen with a 4 hour 2-5A pretreatment, viral replication in cells treated with 2-5A analog I for 20 hours after HCV-PV infection was inhibited by ~70%. The P=S version (analog II) inhibited HCV-PV replication by ~35%. Thus, both 2-5A analogs used here are able to generate an antiviral effect, presumably through RNase L activation. The P=S version, although more resistant to 5' dephosphorylation, did not yield as great an anti-viral effect. It is possible that combination of the 5'-terminal thiophosphate together with the presence of a 3'-inverted abasic moiety can interfere with RNase L activation. Nevertheless, these results demonstrate potent anti-HCV activity by a nuclease-stabilized 2-5A analog.

The level of reduction in HCV-PV replication in cells treated with 2-5A analog I for 20 hours was similar to that in cells pretreated with consensus interferon for 4 hours. To determine if this expanded 2-5A treatment regimen would enhance anti-HCV ribozyme efficacy to the same degree as does the interferon pretreatment, HeLa cells infected with HCV-PV were treated with a combination of 2-5A and anti-HCV ribozyme for 20 hours after infection. In this experiment, a 200 nM treatment with anti-HCV ribozyme or 2-5A treatment alone inhibited viral replication by 88% or ~60%, respectively, compared to SAC treatment (**Figure 38**, left three bars). To maintain consistent transfection conditions but vary the concentration of anti-HCV ribozyme or 2-5A, anti-HCV ribozyme was mixed with the SAC to maintain a total dose of 200 nM. A 50 nM treatment with anti-HCV ribozyme inhibited HCV-PV replication by ~70% (solid middle bar). However, the amount of HCV-PV replication was not further reduced in cells treated with a combination of 50 nM anti-HCV ribozyme and 150 nM 2-5A (striped middle bar). Likewise, cells treated with 100 nM anti-HCV ribozyme inhibited HCV-PV replication by ~80% whether they were also treated with 100 nM of 2-5A or SAC (right two bars). In contrast, antiviral activity increased from 80% to 98% when 100 nM anti-HCV ribozyme was given in combination with interferon (**Figure 36A**). The reasons for the lack of additive or synergistic effects for the ribozyme/2-5A combination therapy is unclear at this time but can be due to that fact that both compounds have a similar mechanism of action (degradation of RNA). Further study is warranted to examine this possibility.

As a monotherapy, 2-5A treatment generates a similar inhibitory effect on HCV-poliovirus replication as does interferon treatment. If these results are maintained in HCV patients, treatment with 2-5A can not only be efficacious but can also generate less side

effects than those observed with interferon if the plethora of interferon-induced genes were not activated.

HBV Cell Culture Models

As previously mentioned, HBV does not infect cells in culture. However, transfection of HBV DNA (either as a head-to-tail dimer or as an "overlength" genome of >100%) into HuH7 or Hep G2 hepatocytes results in viral gene expression and production of HBV virions released into the media. Thus, HBV replication competent DNA are co-transfected with ribozymes in cell culture. Such an approach has been used to report intracellular ribozyme activity against HBV (zu Putlitz, *et al.*, 1999, *J. Virol.*, 73, 5381-5387, and Kim *et al.*, 1999, *Biochem. Biophys. Res. Commun.*, 257, 759-765). In addition, stable hepatocyte cell lines have been generated that express HBV. In these cells, only ribozyme need be delivered; however, performance of a delivery screen is required. Intracellular HBV gene expression can be assayed by a Taqman® assay for HBV RNA or by ELISA for HBV protein. Extracellular virus can be assayed by PCR for DNA or ELISA for protein. Antibodies are commercially available for HBV surface antigen and core protein. A secreted alkaline phosphatase expression plasmid can be used to normalize for differences in transfection efficiency and sample recovery.

HBV Animal Models

There are several small animal models to study HBV replication. One is the transplantation of HBV-infected liver tissue into irradiated mice. Viremia (as evidenced by measuring HBV DNA by PCR) is first detected 8 days after transplantation and peaks between 18 – 25 days (Ilan *et al.*, 1999, *Hepatology*, 29, 553-562).

Transgenic mice that express HBV have also been used as a model to evaluate potential anti-virals. HBV DNA is detectable in both liver and serum (Guidotti *et al.*, 1995, J. Virology, 69, 10, 6158-6169; Morrey *et al.*, 1999, *Antiviral Res.*, 42, 97-108).

An additional model is to establish subcutaneous tumors in nude mice with Hep G2 cells transfected with HBV. Tumors develop in about 2 weeks after inoculation and express HBV surface and core antigens. HBV DNA and surface antigen is also detected in the circulation of tumor-bearing mice (Yao *et al.*, 1996, *J. Viral Hepat.*, 3, 19-22).

In one embodiment, the invention features a mouse, for example a male or female mouse, implanted with HepG2.2.15 cells, wherein the mouse is susceptible to HBV infection and capable of sustaining HBV DNA expression. One embodiment of the invention provides a mouse implanted with HepG2.2.15 cells, wherein said mouse sustains the propagation of

HEPG2.2.15 cells and HBV production (see Macejak, US Provisional Patent Application No. 60/296,876).

Woodchuck hepatitis virus (WHV) is closely related to HBV in its virus structure, genetic organization, and mechanism of replication. As with HBV in humans, persistent WHV infection is common in natural woodchuck populations and is associated with chronic hepatitis and hepatocellular carcinoma (HCC). Experimental studies have established that WHV causes HCC in woodchucks and woodchucks chronically infected with WHV have been used as a model to test a number of anti-viral agents. For example, the nucleoside analogue 3T3 was observed to cause dose dependent reduction in virus (50% reduction after two daily treatments at the highest dose) (Hurwitz *et al.,* 1998. *Antimicrob. Agents Chemother.*, 42, 2804-2809).

<u>HCV Cell Culture Models</u>

Although there have been reports of replication of HCV in cell culture (see below), these systems are difficult to replicate and have proven unreliable. Therefore, as was the case for development of other anti-HCV therapeutics such as interferon and ribavirin, after demonstration of safety in animal studies applicant can proceed directly into a clinical feasibility study.

Several recent reports have documented *in vitro* growth of HCV in human cell lines (Mizutani *et al.,* Biochem Biophys Res Commun 1996 227(3):822-826; Tagawa *et al.,* Journal of Gasteroenterology and Hepatology 1995 10(5):523-527; Cribier *et al., Journal of General Virology* 76(10):2485-2491; Seipp *et al., Journal of General Virology* 1997 78(10)2467-2478; Iacovacci *et al.,* Research Virology 1997 148(2):147-151; Iocavacci *et al.,* Hepatology 1997 26(5) 1328-1337; Ito *et al.,* Journal of General Virology 1996 77(5):1043-1054; Nakajima *et al., Journal of Virology* 1996 70(5):3325-3329; Mizutani *et al., Journal of Virology* 1996 70(10):7219-7223; Valli *et al., Res Virol* 1995 146(4): 285-288; Kato *et al., Biochem Biophys Res Comm* 1995 206(3):863-869). Replication of HCV has been demonstrated in both T and B cell lines as well as cell lines derived from human hepatocytes. Demonstration of replication was documented using either RT-PCR based assays or the b-DNA assay. It is important to note that the most recent publications regarding HCV cell cultures document replication for up to 6-months.

Additionally, another recent study has identified more robust strains of hepatitis C virus having adaptive mutations that allow the strains to replicate more vigorously in human cell culture. The mutations that confer this enhanced ability to replicate are located in a specific region of a protein identified as NS5A. Studies performed at Rockefeller University have shown that in certain cell culture systems, infection with the robust strains produces a 10,000-

fold increase in the number of infected cells. The greatly increased availability of HCV-infected cells in culture can be used to develop high-throughput screening assays, in which a large number of compounds, such as enzymatic nucleic acid molecules, can be tested to determine their effectiveness.

In addition to cell lines that can be infected with HCV, several groups have reported the successful transformation of cell lines with cDNA clones of full-length or partial HCV genomes (Harada *et al.*, Journal of General Virology 1995 76(5)1215-1221; Haramatsu *et al.*, Journal of Viral Hepatitis 1997 4S(1):61-67; Dash *et al.*, American Journal of Pathology 1997 151(2):363-373; Mizuno *et al.*, Gasteroenterology 1995 109(6):1933-40; Yoo *et al.*, Journal Of Virology 1995 69(1):32-38).

HCV Animal Models

The best characterized animal system for HCV infection is the chimpanzee. Moreover, the chronic hepatitis that results from HCV infection in chimpanzees and humans is very similar. Although clinically relevant, the chimpanzee model suffers from several practical impediments that make use of this model difficult. These include; high cost, long incubation requirements and lack of sufficient quantities of animals. Due to these factors, a number of groups have attempted to develop rodent models of chronic hepatitis C infection. While direct infection has not been possible several groups have reported on the stable transfection of either portions or entire HCV genomes into rodents (Yamamoto *et al.*, Hepatology 1995 22(3): 847-855; Galun *et al.*, Journal of Infectious Disease 1995 172(1):25-30; Koike *et al.*, Journal of general Virology 1995 76(12)3031-3038; Pasquinelli *et al.*, Hepatology 1997 25(3): 719-727; Hayashi *et al.*, Princess Takamatsu Symp 1995 25:1430149; Mariya K, Yotsuyanagi H, Shintani Y, Fujie H, Ishibashi K, Matsuura Y, Miyamura T, Koike K. Hepatitis C virus core protein induces hepatic steatosis in transgenic mice. Journal of General Virology 1997 78(7) 1527-1531; Takehara *et al.*, Hepatology 1995 21(3):746-751; Kawamura *et al.*, Hepatology 1997 25(4): 1014-1021). In addition, transplantation of HCV infected human liver into immunocompromised mice results in prolonged detection of HCV RNA in the animal's blood.

Vierling, International PCT Publication No. WO 99/16307, describes a method for expressing hepatitis C virus in an *in vivo* animal model. Viable, HCV infected human hepatocytes are transplanted into a liver parenchyma of a scid/scid mouse host. The scid/scid mouse host is then maintained in a viable state, whereby viable, morphologically intact human hepatocytes persist in the donor tissue and hepatitis C virus is replicated in the persisting human hepatocytes. This model provides an effective means for the study of HCV inhibition by enzymatic nucleic acids *in vivo*.

Indications

Particular degenerative and disease states that can be associated with HBV expression modulation include, but are not limited to, HBV infection, hepatitis, cancer, tumorigenesis, cirrhosis, liver failure and other conditions related to the level of HBV.

Particular degenerative and disease states that can be associated with HCV expression modulation include, but are not limited to, HCV infection, hepatitis, cancer, tumorigenesis, cirrhosis, liver failure and other conditions related to the level of HCV.

The present body of knowledge in HBV and HCV research indicates the need for methods to assay HBV or HCV activity and for compounds that can regulate HBV and HCV expression for research, diagnostic, and therapeutic use.

Lamivudine (3TC®), L-FMAU, adefovir dipivoxil, type 1 Interferon (*e.g*, interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon 2b, and polyethylene glycol consensus interferon), therapeutic vaccines, steriods, and 2'-5' Oligoadenylates are non-limiting examples of pharmaceutical agents that can be combined with or used in conjunction with the nucleic acid molecules (*e.g.* ribozymes and antisense molecules) of the instant invention. Those skilled in the art will recognize that other drugs or other therapies can similarly and readily be combined with the nucleic acid molecules of the instant invention (*e.g.* ribozymes and antisense molecules) and are, therefore, within the scope of the instant invention.

Diagnostic uses

The nucleic acid molecules of this invention can be used as diagnostic tools to examine genetic drift and mutations within diseased cells or to detect the presence of HBV or HCV RNA in a cell. For example, the close relationship between enzymatic nucleic acid activity and the structure of the target RNA allows the detection of mutations in any region of the molecule which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple enzymatic nucleic acids described in this invention, one can map nucleotide changes which are important to RNA structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with enzymatic nucleic acids can be used to inhibit gene expression and define the role (essentially) of specified gene products in the progression of disease. In this manner, other genetic targets can be defined as important mediators of the disease. These experiments can lead to better treatment of the disease progression by affording the possibility of combinational therapies (*e.g.*, multiple enzymatic nucleic acid molecules targeted to different genes, enzymatic nucleic acid molecules coupled

with known small molecule inhibitors, or intermittent treatment with combinations of enzymatic nucleic acid molecules and/or other chemical or biological molecules). Other *in vitro* uses of enzymatic nucleic acid moleculesof this invention are well known in the art, and include detection of the presence of mRNAs associated with HBV or HCV-related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with an enzymatic nucleic acid using standard methodology.

In a specific example, enzymatic nucleic acid molecules which can cleave only wild-type or mutant forms of the target RNA are used for the assay. The first enzymatic nucleic acid is used to identify wild-type RNA present in the sample and the second enzymatic nucleic acid is used to identify mutant RNA in the sample. As reaction controls, synthetic substrates of both wild-type and mutant RNA can be cleaved by both enzymatic nucleic acid molecules to demonstrate the relative ribozyme efficiencies in the reactions and the absence of cleavage of the "non-targeted" RNA species. The cleavage products from the synthetic substrates can also serve to generate size markers for the analysis of wild-type and mutant RNAs in the sample population. Thus each analysis involves two enzymatic nucleic acid molecules, two substrates and one unknown sample which is combined into six reactions. The presence of cleavage products is determined using an RNAse protection assay so that full-length and cleavage fragments of each RNA can be analyzed in one lane of a polyacrylamide gel. It is not absolutely required to quantify the results to gain insight into the expression of mutant RNAs and putative risk of the desired phenotypic changes in target cells. The expression of mRNA whose protein product is implicated in the development of the phenotype (*i.e.*, HBV or HCV) is adequate to establish risk. If probes of comparable specific activity are used for both transcripts, then a qualitative comparison of RNA levels is adequate and will decrease the cost of the initial diagnosis. Higher mutant form to wild-type ratios are correlated with higher risk whether RNA levels are compared qualitatively or quantitatively.

Additional Uses

Potential usefulness of sequence-specific enzymatic nucleic acid molecules of the instant invention have many of the same applications for the study of RNA that DNA restriction endonucleases have for the study of DNA (Nathans *et al.*, 1975 *Ann. Rev. Biochem.* 44:273). For example, the pattern of restriction fragments can be used to establish sequence relationships between two related RNAs, and large RNAs can be specifically cleaved to fragments of a size more useful for study. The ability to engineer sequence specificity of the enzymatic nucleic acid molecule is ideal for cleavage of RNAs of unknown sequence. Applicant describes the use of nucleic acid molecules to down-regulate gene

expression of target genes in bacterial, microbial, fungal, viral, and eukaryotic systems including plant, or mammalian cells.

All patents and publications mentioned in the specification are indicative of the levels of skill of those skilled in the art to which the invention pertains. All references cited in this disclosure are incorporated by reference to the same extent as if each reference had been incorporated by reference in its entirety individually.

One skilled in the art would readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The methods and compositions described herein as presently representative of preferred embodiments are exemplary and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art, which are encompassed within the spirit of the invention, are defined by the scope of the claims.

It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention. Thus, such additional embodiments are within the scope of the present invention and the following claims.

The invention illustratively described herein suitably can be practiced in the absence of any element or elements, limitation or limitations that are not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising", "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments, optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the description and the appended claims.

In addition, where features or aspects of the invention are described in terms of Markush groups or other grouping of alternatives, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group or other group.

TABLE I

Characteristics of naturally occurring ribozymes

Group I Introns

- Size: ~150 to >1000 nucleotides.
- Requires a U in the target sequence immediately 5' of the cleavage site.
- Binds 4-6 nucleotides at the 5'-side of the cleavage site.
- Reaction mechanism: attack by the 3'-OH of guanosine to generate cleavage products with 3'-OH and 5'-guanosine.
- Additional protein cofactors required in some cases to help folding and maintainance of the active structure.
- Over 300 known members of this class. Found as an intervening sequence in *Tetrahymena thermophila* rRNA, fungal mitochondria, chloroplasts, phage T4, blue-green algae, and others.
- Major structural features largely established through phylogenetic comparisons, mutagenesis, and biochemical studies [i,ii].
- Complete kinetic framework established for one ribozyme [iii,iv,v,vi].
- Studies of ribozyme folding and substrate docking underway [vii,viii,ix].
- Chemical modification investigation of important residues well established [x,xi].
- The small (4-6 nt) binding site may make this ribozyme too non-specific for targeted RNA cleavage, however, the Tetrahymena group I intron has been used to repair a "defective" β-galactosidase message by the ligation of new β-galactosidase sequences onto the defective message [xii].

RNAse P RNA (M1 RNA)

- Size: ~290 to 400 nucleotides.
- RNA portion of a ubiquitous ribonucleoprotein enzyme.

- Cleaves tRNA precursors to form mature tRNA [xiii].
- Reaction mechanism: possible attack by M^{2+}-OH to generate cleavage products with 3′-OH and 5′-phosphate.
- RNAse P is found throughout the prokaryotes and eukaryotes. The RNA subunit has been sequenced from bacteria, yeast, rodents, and primates.
- Recruitment of endogenous RNAse P for therapeutic applications is possible through hybridization of an External Guide Sequence (EGS) to the target RNA [xiv,xv]
- Important phosphate and 2′ OH contacts recently identified [xvi,xvii]

Group II Introns

- Size: >1000 nucleotides.
- Trans cleavage of target RNAs recently demonstrated [xviii,xix].
- Sequence requirements not fully determined.
- Reaction mechanism: 2′-OH of an internal adenosine generates cleavage products with 3′-OH and a "lariat" RNA containing a 3′-5′ and a 2′-5′ branch point.
- Only natural ribozyme with demonstrated participation in DNA cleavage [xx,xxi] in addition to RNA cleavage and ligation.
- Major structural features largely established through phylogenetic comparisons [xxii].
- Important 2′ OH contacts beginning to be identified [xxiii]
- Kinetic framework under development [xxiv]

Neurospora VS RNA

- Size: ~144 nucleotides.
- Trans cleavage of hairpin target RNAs recently demonstrated [xxv].

- Sequence requirements not fully determined.
- Reaction mechanism: attack by 2'-OH 5' to the scissile bond to generate cleavage products with 2',3'-cyclic phosphate and 5'-OH ends.
- Binding sites and structural requirements not fully determined.
- Only 1 known member of this class. Found in Neurospora VS RNA.

Hammerhead Ribozyme
(see text for references)

- Size: ~13 to 40 nucleotides.
- Requires the target sequence UH immediately 5' of the cleavage site.
- Binds a variable number nucleotides on both sides of the cleavage site.
- Reaction mechanism: attack by 2'-OH 5' to the scissile bond to generate cleavage products with 2',3'-cyclic phosphate and 5'-OH ends.
- 14 known members of this class. Found in a number of plant pathogens (virusoids) that use RNA as the infectious agent.
- Essential structural features largely defined, including 2 crystal structures [[xxvi,xxvii]]
- Minimal ligation activity demonstrated (for engineering through *in vitro* selection) [[xxviii]]
- Complete kinetic framework established for two or more ribozymes [[xxix]].
- Chemical modification investigation of important residues well established [[xxx]].

Hairpin Ribozyme

- Size: ~50 nucleotides.
- Requires the target sequence GUC immediately 3' of the cleavage site.

- Binds 4-6 nucleotides at the 5'-side of the cleavage site and a variable number to the 3'-side of the cleavage site.
- Reaction mechanism: attack by 2'-OH 5' to the scissile bond to generate cleavage products with 2',3'-cyclic phosphate and 5'-OH ends.
- 3 known members of this class. Found in three plant pathogen (satellite RNAs of the tobacco ringspot virus, arabis mosaic virus and chicory yellow mottle virus) which uses RNA as the infectious agent.
- Essential structural features largely defined [xxxi,xxxii,xxxiii,xxxiv]
- Ligation activity (in addition to cleavage activity) makes ribozyme amenable to engineering through *in vitro* selection [xxxv]
- Complete kinetic framework established for one ribozyme [xxxvi].
- Chemical modification investigation of important residues begun [xxxvii,xxxviii].

Hepatitis Delta Virus (HDV) Ribozyme

- Size: ~60 nucleotides.
- Trans cleavage of target RNAs demonstrated [xxxix].
- Binding sites and structural requirements not fully determined, although no sequences 5' of cleavage site are required. Folded ribozyme contains a pseudoknot structure [xl].
- Reaction mechanism: attack by 2'-OH 5' to the scissile bond to generate cleavage products with 2',3'-cyclic phosphate and 5'-OH ends.
- Only 2 known members of this class. Found in human HDV.
- [xli]Circular form of HDV is active and shows increased nuclease stability [xlii]

---

i. Michel, Francois; Westhof, Eric. Slippery substrates. Nat. Struct. Biol. (1994), 1(1), 5-7.

ii. Lisacek, Frederique; Diaz, Yolande; Michel, Francois. Automatic identification of group I intron cores in genomic DNA sequences. J. Mol. Biol. (1994), 235(4), 1206-17.

iii. Herschlag, Daniel; Cech, Thomas R.. Catalysis of RNA cleavage by the Tetrahymena thermophila ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. Biochemistry (1990), 29(44), 10159-71.

[iv]. Herschlag, Daniel; Cech, Thomas R.. Catalysis of RNA cleavage by the Tetrahymena thermophila ribozyme. 2. Kinetic description of the reaction of an RNA substrate that forms a mismatch at the active site. Biochemistry (1990), 29(44), 10172-80.

[v]. Knitt, Deborah S.; Herschlag, Daniel. pH Dependencies of the Tetrahymena Ribozyme Reveal an Unconventional Origin of an Apparent pKa. Biochemistry (1996), 35(5), 1560-70.

[vi]. Bevilacqua, Philip C.; Sugimoto, Naoki; Turner, Douglas H.. A mechanistic framework for the second step of splicing catalyzed by the Tetrahymena ribozyme. Biochemistry (1996), 35(2), 648-58.

[vii]. Li, Yi; Bevilacqua, Philip C.; Mathews, David; Turner, Douglas H.. Thermodynamic and activation parameters for binding of a pyrene-labeled substrate by the Tetrahymena ribozyme: docking is not diffusion-controlled and is driven by a favorable entropy change. Biochemistry (1995), 34(44), 14394-9.

[viii]. Banerjee, Aloke Raj; Turner, Douglas H.. The time dependence of chemical modification reveals slow steps in the folding of a group I ribozyme. Biochemistry (1995), 34(19), 6504-12.

[ix]. Zarrinkar, Patrick P.; Williamson, James R.. The P9.1-P9.2 peripheral extension helps guide folding of the Tetrahymena ribozyme. Nucleic Acids Res. (1996), 24(5), 854-8.

[x]. Strobel, Scott A.; Cech, Thomas R.. Minor groove recognition of the conserved G.cntdot.U pair at the Tetrahymena ribozyme reaction site. Science (Washington, D. C.) (1995), 267(5198), 675-9.

[xi]. Strobel, Scott A.; Cech, Thomas R.. Exocyclic Amine of the Conserved G.cntdot.U Pair at the Cleavage Site of the Tetrahymena Ribozyme Contributes to 5'-Splice Site Selection and Transition State Stabilization. Biochemistry (1996), 35(4), 1201-11.

[xii]. Sullenger, Bruce A.; Cech, Thomas R.. Ribozyme-mediated repair of defective mRNA by targeted trans-splicing. Nature (London) (1994), 371(6498), 619-22.

[xiii]. Robertson, H.D.; Altman, S.; Smith, J.D. J. Biol. Chem., 247, 5243-5251 (1972).

[xiv]. Forster, Anthony C.; Altman, Sidney. External guide sequences for an RNA enzyme. Science (Washington, D. C., 1883-) (1990), 249(4970), 783-6.

[xv]. Yuan, Y.; Hwang, E. S.; Altman, S. Targeted cleavage of mRNA by human RNase P. Proc. Natl. Acad. Sci. USA (1992) 89, 8006-10.

[xvi]. Harris, Michael E.; Pace, Norman R.. Identification of phosphates involved in catalysis by the ribozyme RNase P RNA. RNA (1995), 1(2), 210-18.

[xvii]. Pan, Tao; Loria, Andrew; Zhong, Kun. Probing of tertiary interactions in RNA: 2'-hydroxyl-base contacts between the RNase P RNA and pre-tRNA. Proc. Natl. Acad. Sci. U. S. A. (1995), 92(26), 12510-14.

[xviii]. Pyle, Anna Marie; Green, Justin B.. Building a Kinetic Framework for Group II Intron Ribozyme Activity: Quantitation of Interdomain Binding and Reaction Rate. Biochemistry (1994), 33(9), 2716-25.

[xix]. Michels, William J. Jr.; Pyle, Anna Marie. Conversion of a Group II Intron into a New Multiple-Turnover Ribozyme that Selectively Cleaves Oligonucleotides: Elucidation of Reaction Mechanism and Structure/Function Relationships. Biochemistry (1995), 34(9), 2965-77.

[xx]. Zimmerly, Steven; Guo, Huatao; Eskes, Robert; Yang, Jian; Perlman, Philip S.; Lambowitz, Alan M.. A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. Cell (Cambridge, Mass.) (1995), 83(4), 529-38.

[xxi]. Griffin, Edmund A., Jr.; Qin, Zhifeng; Michels, Williams J., Jr.; Pyle, Anna Marie. Group II intron ribozymes that cleave DNA and RNA linkages with similar efficiency, and lack contacts with substrate 2'-hydroxyl groups. Chem. Biol. (1995), 2(11), 761-70.

[xxii]. Michel, Francois; Ferat, Jean Luc. Structure and activities of group II introns. Annu. Rev. Biochem. (1995), 64, 435-61.

[xxiii]. Abramovitz, Dana L.; Friedman, Richard A.; Pyle, Anna Marie. Catalytic role of 2'-hydroxyl groups within a group II intron active site. Science (Washington, D. C.) (1996), 271(5254), 1410-13.

[xxiv]. Daniels, Danette L.; Michels, William J., Jr.; Pyle, Anna Marie. Two competing pathways for self-splicing by group II introns: a quantitative analysis of in vitro reaction rates and products. J. Mol. Biol. (1996), 256(1), 31-49.

[xxv]. Guo, Hans C. T.; Collins, Richard A.. Efficient trans-cleavage of a stem-loop RNA substrate by a ribozyme derived from Neurospora VS RNA. EMBO J. (1995), 14(2), 368-76.

[xxvi]. Scott, W.G., Finch, J.T., Aaron,K. The crystal structure of an all RNA hammerhead ribozyme:Aproposed mechanism for RNA catalytic cleavage. Cell, (1995), 81, 991-1002.

[xxvii]. McKay, Structure and function of the hammerhead ribozyme: an unfinished story. RNA, (1996), 2, 395-403.
[xxviii]. Long, D., Uhlenbeck, O., Hertel, K. Ligation with hammerhead ribozymes. US Patent No. 5,633,133.
[xxix]. Hertel, K.J., Herschlag, D., Uhlenbeck, O. A kinetic and thermodynamic framework for the hammerhead ribozyme reaction. Biochemistry, (1994) 33, 3374-3385.Beigelman, L., *et al.*, Chemical modifications of hammerhead ribozymes. J. Biol. Chem., (1995) 270, 25702-25708.
[xxx]. Beigelman, L., *et al.*, Chemical modifications of hammerhead ribozymes. J. Biol. Chem., (1995) 270, 25702-25708.
[xxxi]. Hampel, Arnold; Tritz, Richard; Hicks, Margaret; Cruz, Phillip. 'Hairpin' catalytic RNA model: evidence for helixes and sequence requirement for substrate RNA. Nucleic Acids Res. (1990), 18(2), 299-304.
[xxxii]. Chowrira, Bharat M.; Berzal-Herranz, Alfredo; Burke, John M.. Novel guanosine requirement for catalysis by the hairpin ribozyme. Nature (London) (1991), 354(6351), 320-2.
[xxxiii]. Berzal-Herranz, Alfredo; Joseph, Simpson; Chowrira, Bharat M.; Butcher, Samuel E.; Burke, John M.. Essential nucleotide sequences and secondary structure elements of the hairpin ribozyme. EMBO J. (1993), 12(6), 2567-73.
[xxxiv]. Joseph, Simpson; Berzal-Herranz, Alfredo; Chowrira, Bharat M.; Butcher, Samuel E.. Substrate selection rules for the hairpin ribozyme determined by in vitro selection, mutation, and analysis of mismatched substrates. Genes Dev. (1993), 7(1), 130-8.
[xxxv]. Berzal-Herranz, Alfredo; Joseph, Simpson; Burke, John M.. In vitro selection of active hairpin ribozymes by sequential RNA-catalyzed cleavage and ligation reactions. Genes Dev. (1992), 6(1), 129-34.
[xxxvi]. Hegg, Lisa A.; Fedor, Martha J.. Kinetics and Thermodynamics of Intermolecular Catalysis by Hairpin Ribozymes. Biochemistry (1995), 34(48), 15813-28.
[xxxvii]. Grasby, Jane A.; Mersmann, Karin; Singh, Mohinder; Gait, Michael J.. Purine Functional Groups in Essential Residues of the Hairpin Ribozyme Required for Catalytic Cleavage of RNA. Biochemistry (1995), 34(12), 4068-76.
[xxxviii]. Schmidt, Sabine; Beigelman, Leonid; Karpeisky, Alexander; Usman, Nassim; Sorensen, Ulrik S.; Gait, Michael J.. Base and sugar requirements for RNA cleavage of essential nucleoside residues in internal loop B of the hairpin ribozyme: implications for secondary structure. Nucleic Acids Res. (1996), 24(4), 573-81.
[xxxix]. Perrotta, Anne T.; Been, Michael D.. Cleavage of oligoribonucleotides by a ribozyme derived from the hepatitis .delta. virus RNA sequence. Biochemistry (1992), 31(1), 16-21.
[xl]. Perrotta, Anne T.; Been, Michael D.. A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. Nature (London) (1991), 350(6317), 434-6.
[xli]
[xlii]. Puttaraju, M.; Perrotta, Anne T.; Been, Michael D.. A circular trans-acting hepatitis delta virus ribozyme. Nucleic Acids Res. (1993), 21(18), 4253-8.

Table II:

A. 2.5 µmol Synthesis Cycle ABI 394 Instrument

Reagent	Equivalents	Amount	Wait Time* DNA	Wait Time* 2'-O-methyl	Wait Time*RNA
Phosphoramidites	6.5	163 µL	45 sec	2.5 min	7.5 min
S-Ethyl Tetrazole	23.8	238 µL	45 sec	2.5 min	7.5 min
Acetic Anhydride	100	233 µL	5 sec	5 sec	5 sec
N-Methyl Imidazole	186	233 µL	5 sec	5 sec	5 sec
TCA	176	2.3 mL	21 sec	21 sec	21 sec
Iodine	11.2	1.7 mL	45 sec	45 sec	45 sec
Beaucage	12.9	645 µL	100 sec	300 sec	300 sec
Acetonitrile	NA	6.67 mL	NA	NA	NA

B. 0.2 µmol Synthesis Cycle ABI 394 Instrument

Reagent	Equivalents	Amount	Wait Time* DNA	Wait Time* 2'-O-methyl	Wait Time*RNA
Phosphoramidites	15	31 µL	45 sec	233 sec	465 sec
S-Ethyl Tetrazole	38.7	31 µL	45 sec	233 min	465 sec
Acetic Anhydride	655	124 µL	5 sec	5 sec	5 sec
N-Methyl Imidazole	1245	124 µL	5 sec	5 sec	5 sec
TCA	700	732 µL	10 sec	10 sec	10 sec
Iodine	20.6	244 µL	15 sec	15 sec	15 sec
Beaucage	7.7	232 µL	100 sec	300 sec	300 sec
Acetonitrile	NA	2.64 mL	NA	NA	NA

C. 0.2 µmol Synthesis Cycle 96 well Instrument

Reagent	Equivalents:DNA/ 2'-O-methyl/Ribo	Amount: DNA/2'-O-methyl/Ribo	Wait Time* DNA	Wait Time* 2'-O-methyl	Wait Time* Ribo
Phosphoramidites	22/33/66	40/60/120 µL	60 sec	180 sec	360sec
S-Ethyl Tetrazole	70/105/210	40/60/120 µL	60 sec	180 min	360 sec
Acetic Anhydride	265/265/265	50/50/50 µL	10 sec	10 sec	10 sec
N-Methyl Imidazole	502/502/502	50/50/50 µL	10 sec	10 sec	10 sec
TCA	238/475/475	250/500/500 µL	15 sec	15 sec	15 sec
Iodine	6.8/6.8/6.8	80/80/80 µL	30 sec	30 sec	30 sec
Beaucage	34/51/51	80/120/120	100 sec	200 sec	200 sec
Acetonitrile	NA	1150/1150/1150 µL	NA	NA	NA

- Wait time does not include contact time during delivery.

Table III: HBV Strains and Accession numbers

Accession Number	NAME
AF100308.1	AF100308 Hepatitis B virus strain 2-18, complete
AB026815.1	AB026815 Hepatitis B virus DNA, complete genome,
AB033559.1	AB033559 Hepatitis B virus DNA, complete genome,
AB033558.1	AB033558 Hepatitis B virus DNA, complete genome,
AB033557.1	AB033557 Hepatitis B virus DNA, complete genome,
AB033556.1	AB033556 Hepatitis B virus DNA, complete genome,
AB033555.1	AB033555 Hepatitis B virus DNA, complete genome,
AB033554.1	AB033554 Hepatitis B virus DNA, complete genome,
AB033553.1	AB033553 Hepatitis B virus DNA, complete genome,
AB033552.1	AB033552 Hepatitis B virus DNA, complete genome,
AB033551.1	AB033551 Hepatitis B virus DNA, complete genome,
AB033550.1	AB033550 Hepatitis B virus DNA, complete genome
AF143308.1	AF143308 Hepatitis B virus clone WB1254, complete
AF143307.1	AF143307 Hepatitis B virus clone RM518, complete
AF143306.1	AF143306 Hepatitis B virus clone RM517, complete
AF143305.1	AF143305 Hepatitis B virus clone RM501, complete
AF143304.1	AF143304 Hepatitis B virus clone HD319, complete
AF143303.1	AF143303 Hepatitis B virus clone HD1406, complete
AF143302.1	AF143302 Hepatitis B virus clone HD1402, complete
AF143301.1	AF143301 Hepatitis B virus clone BW1903, complete
AF143300.1	AF143300 Hepatitis B virus clone 7832-G4, complete
AF143299.1	AF143299 Hepatitis B virus clone 7744-G9, complete
AF143298.1	AF143298 Hepatitis B virus clone 7720-G8, complete
AB026814.1	AB026814 Hepatitis B virus DNA, complete genome,
AB026813.1	AB026813 Hepatitis B virus DNA, complete genome,
AB026812.1	AB026812 Hepatitis B virus DNA, complete genome,
AB026811.1	AB026811 Hepatitis B virus DNA, complete genome,
AJ131956.1	HBV131956 Hepatitis B virus complete genome,
AF151735.1	AF151735 Hepatitis B virus, complete genome
AF090842.1	AF090842 Hepatitis B virus strain G5.27295, complete
AF090841.1	AF090841 Hepatitis B virus strain G4.27241, complete
AF090840.1	AF090840 Hepatitis B virus strain G3.27270, complete
AF090839.1	AF090839 Hepatitis B virus strain G2.27246, complete
AF090838.1	AF090838 Hepatitis B virus strain P1.27239, complete
Y18858.1	HBV18858 Hepatitis B virus complete genome, isolate
Y18857.1	HBV18857 Hepatitis B virus complete genome, isolate
D12980.1	HPBCG Hepatitis B virus subtype adr(SRADR) DNA,
Y18856.1	HBV18856 Hepatitis B virus complete genome, isolate
Y18855.1	HBV18855 Hepatitis B virus complete genome, isolate
AJ131133.1	HBV131133 Hepatitis B virus, complete genome, strain
X80925.1	HBVP6PCXX Hepatitis B virus (patient 6) complete
X80926.1	HBVP5PCXX Hepatitis B virus (patient 5) complete
X80924.1	HBVP4PCXX Hepatitis B virus (patient 4) complete

AF100309.1	Hepatitis B virus strain 56, complete genome
AF068756.1	AF068756 Hepatitis B virus, complete genome
AF043593.1	AF043593 Hepatitis B virus isolate 6/89, complete
Y07587.1	HBVAYWGEN Hepatitis B virus, complete genome
D28880.1	D28880 Hepatitis B virus DNA, complete genome, strain
X98076.1	HBVDEFVP3 Hepatitis B virus complete genome with
X98075.1	HBVDEFVP2 Hepatitis B virus complete genome with
X98074.1	HBVDEFVP1 Hepatitis B virus complete genome with
X98077.1	HBVCGWITY Hepatitis B virus complete genome, wild type
X98072.1	HBVCGINSC Hepatitis B virus complete genome with
X98073.1	HBVCGINCX Hepatitis B virus complete genome with
U95551.1	U95551 Hepatitis B virus subtype ayw, complete genome
D23684.1	HPBC6T588 Hepatitis B virus (C6-TKB588) complete genome
D23683.1	HPBC5HKO2 Hepatitis B virus (C5-HBVKO2) complete genome
D23682.1	HPBB5HKO1 Hepatitis B virus (B5-HBVKO1) complete genome
D23681.1	HPBC4HST2 Hepatitis B virus (C4-HBVST2) complete genome
D23680.1	HPBB4HST1 Hepatitis B virus (B4-HBVST1) complete genome
D00331.1	HPBADW3 Hepatitis B virus genome, complete genome
D00330.1	HPBADW2 Hepatitis B virus genome, complete genome
D50489.1	HPBA11A Hepatitis B virus DNA, complete genome
D23679.1	HPBA3HMS2 Hepatitis B virus (A3-HBVMS2) complete genome
D23678.1	HPBA2HYS2 Hepatitis B virus (A2-HBVYS2) complete genome
D23677.1	HPBA1HKK2 Hepatitis B virus (A1-HBVKK2) complete genome
D16665.1	HPBADRM Hepatitis B virus DNA, complete genome
D00329.1	HPBADW1 Hepatitis B virus (HBV) genome, complete genome
X97851.1	HBVP6CSX Hepatitis B virus (patient 6) complete genome
X97850.1	HBVP4CSX Hepatitis B virus (patient 4) complete genome
X97849.1	HBVP3CSX Hepatitis B virus (patient 3) complete genome
X97848.1	HBVP2CSX Hepatitis B virus (patient 2) complete genome
X51970.1	HVHEPB Hepatitis B virus (HBV 991) complete genome
M38636.1	HPBCGADR Hepatitis B virus, subtype adr, complete genome
X59795.1	HBVAYWMCG Hepatitis B virus (ayw subtype mutant)
M38454.1	HPBADR1CG Hepatitis B virus , complete genome
M32138.1	HPBHBVAA Hepatitis B virus variant HBV-alpha1, complete
J02203.1	HPBAYW Human hepatitis B virus (subtype ayw), complete
M12906.1	HPBADRA Hepatitis B virus subtype adr, complete genome
M54923.1	HPBADWZ Hepatitis B virus (subtype adw), complete genome
L27106.1	HPBMUT Hepatitis B virus mutant complete genome

Table IV: HBV Substrate Sequence

NT Position*	SUBSTRATE	SEQ ID
82	CUAUCGUCCCCUUCUUCAUC	1.
101	CUACCGUUCCGGCC	2.
159	CUUCUCAUCU	3.
184	CUUCCCUUCACCAC	4.
269	GACUCUCAGAAUGUCAACGAC	5.
381	CUGUAGGCAUAAAUGGUCUG	6.
401	GUUCACCAGCACCAUGCAACUUUUU	7.
424	UUUCACGUCUGCCUAAUCAUC	8.
524	AUUUGGAGCUUC	9.
562	CUGACUUCUUUCCUUCUAUUC	10.
649	CUCACCAUACCGCACUCA	11.
667	GGCAAGCUAUUCUGUG	12.
717	GGAAGUAAUUUGGAAGAC	13.
758	CAGCUAUGUCAAUGUUAA	14.
783	CUAAAAUCGGCCUAAAAUCAGAC	15.
812	CAUUUCCUGUCUCACUUUUGGAAGAG	16.
887	UCCUGCUUACAGAC	17.
922	CAACACUUCCGGAAACUACUGUUGUUAG	18.
989	CUCGCCUCGCAGACGAAGGUCUC	19.
1009	CAAUCGCCGCGUCGCAGAAG	20.
1031	AUCUCAAUCUCGGGAAUCUCAA	21.
1052	AUGUUAGUAUCCCUUGGACUC	22.
1072	CAUAAGGUGGGAAACUUUACUG	23.
1109	CUGUACCUAUUCUUUAAAUCC	24.
1127	CUGAGUGGCAAACUCCC	25.
1271	CCAAAUAUCUGCCCUUGGACAA	26.
1297	AUUAAACCAUAUUAUCCUGAACA	27.
1319	AUGCAGUUAAUCAUUACUUCAAAACUA	28.
1340	AAACUAGGCAUUA	29.
1370	AGGCGGGCAUUCUAUAUAAGAGAG	30.
1393	GAAACUACGCGCAGCGCCUCAUUUUGU	31.
1412	CAUUUUGUGGGUCACCAUA	32.
1441	CAAGAGCUACAGCAUGGG	33.

LOCUS HPBADR1CG 3221 bp DNA circular VRL
06-MAR-1995
DEFINITION Hepatitis B virus , complete genome.
ACCESSION M38454

*The nucleotide number referred to in that table is the position of the 5' end of the oligo in this sequence.

TABLE V: HUMAN HBV HAMMERHEAD RIBOZYME AND TARGET SEQUENCE

Pos	Substrate	Seq ID	Hammerhead	Seq ID
13	CCACCACU U UCCACCAA	34	UUGGUGGA CUGAUGAG GCCGUUAGGC CGAA AGUGGUGG	7434
14	CACCACUU U CCACCAAA	35	UUUGGUGG CUGAUGAG GCCGUUAGGC CGAA AAGUGGUG	7435
15	ACCACUUU C CACCAAAC	36	GUUUGGUG CUGAUGAG GCCGUUAGGC CGAA AAAGUGGU	7436
25	ACCAAACU C UUCAAGAU	37	AUCUUGAA CUGAUGAG GCCGUUAGGC CGAA AGUUUGGU	7437
27	CAAACUCU U CAAGAUCC	38	GGAUCUUG CUGAUGAG GCCGUUAGGC CGAA AGAGUUUG	7438
28	AAACUCUU C AAGAUCCC	39	GGGAUCUU CUGAUGAG GCCGUUAGGC CGAA AAGAGUUU	7439
34	UUCAAGAU C CCAGAGUC	40	GACUCUGG CUGAUGAG GCCGUUAGGC CGAA AUCUUGAA	7440
42	CCCAGAGU C AGGGCCCU	41	AGGGCCCU CUGAUGAG GCCGUUAGGC CGAA ACUCUGGG	7441
53	GGCCCUGU A CUUUCCUG	42	CAGGAAAG CUGAUGAG GCCGUUAGGC CGAA ACAGGGCC	7442
56	CCUGUACU U UCCUGCUG	43	CAGCAGGA CUGAUGAG GCCGUUAGGC CGAA AGUACAGG	7443
57	CUGUACUU U CCUGCUGG	44	CCAGCAGG CUGAUGAG GCCGUUAGGC CGAA AAGUACAG	7444
58	UGUACUUU C CUGCUGGU	45	ACCAGCAG CUGAUGAG GCCGUUAGGC CGAA AAAGUACA	7445
71	UGGUGGCU C CAGUUCAG	46	CUGAACUG CUGAUGAG GCCGUUAGGC CGAA AGCCACCA	7446
76	GCUCCAGU U CAGGAACA	47	UGUUCCUG CUGAUGAG GCCGUUAGGC CGAA ACUGGAGC	7447
77	CUCCAGUU C AGGAACAG	48	CUGUUCCU CUGAUGAG GCCGUUAGGC CGAA AACUGGAG	7448
97	GCCCUGCU C AGAAUACU	49	AGUAUUCU CUGAUGAG GCCGUUAGGC CGAA AGCAGGGC	7449
103	CUCAGAAU A CUGUCUCU	50	AGAGACAG CUGAUGAG GCCGUUAGGC CGAA AUUCUGAG	7450
108	AAUACUGU C UCUGCCAU	51	AUGGCAGA CUGAUGAG GCCGUUAGGC CGAA ACAGUAUU	7451
110	UACUGUCU C UGCCAUAU	52	AUAUGGCA CUGAUGAG GCCGUUAGGC CGAA AGACAGUA	7452
117	UCUGCCAU A UCGUCAAU	53	AUUGACGA CUGAUGAG GCCGUUAGGC CGAA AUGGCAGA	7453
119	UGCCAUAU C GUCAAUCU	54	AGAUUGAC CUGAUGAG GCCGUUAGGC CGAA AUAUGGCA	7454
122	CAUAUCGU C AAUCUUAU	55	AUAAGAUU CUGAUGAG GCCGUUAGGC CGAA ACGAUAUG	7455
126	UCGUCAAU C UUAUCGAA	56	UUCGAUAA CUGAUGAG GCCGUUAGGC CGAA AUUGACGA	7456
128	GUCAAUCU U AUCGAAGA	57	UCUUCGAU CUGAUGAG GCCGUUAGGC CGAA AGAUUGAC	7457
129	UCAAUCUU A UCGAAGAC	58	GUCUUCGA CUGAUGAG GCCGUUAGGC CGAA AAGAUUGA	7458
131	AAUCUUAU C GAAGACUG	59	CAGUCUUC CUGAUGAG GCCGUUAGGC CGAA AUAAGAUU	7459
150	GACCCUGU A CCGAACAU	60	AUGUUCGG CUGAUGAG GCCGUUAGGC CGAA ACAGGGUC	7460
168	GAGAACAU C GCAUCAGG	61	CCUGAUGC CUGAUGAG GCCGUUAGGC CGAA AUGUUCUC	7461
173	CAUCGCAU C AGGACUCC	62	GGAGUCCU CUGAUGAG GCCGUUAGGC CGAA AUGCGAUG	7462
180	UCAGGACU C CUAGGACC	63	GGUCCUAG CUGAUGAG GCCGUUAGGC CGAA AGUCCUGA	7463
183	GGACUCCU A GGACCCCU	64	AGGGGUCC CUGAUGAG GCCGUUAGGC CGAA AGGAGUCC	7464
195	CCCCUGCU C GUGUUACA	65	UGUAACAC CUGAUGAG GCCGUUAGGC CGAA AGCAGGGG	7465
200	GCUCGUGU U ACAGGCGG	66	CCGCCUGU CUGAUGAG GCCGUUAGGC CGAA ACACGAGC	7466
201	CUCGUGUU A CAGGCGGG	67	CCCGCCUG CUGAUGAG GCCGUUAGGC CGAA AACACGAG	7467
212	GGCGGGGU U UUUCUUGU	68	ACAAGAAA CUGAUGAG GCCGUUAGGC CGAA ACCCCGCC	7468
213	GCGGGGUU U UUCUUGUU	69	AACAAGAA CUGAUGAG GCCGUUAGGC CGAA AACCCCGC	7469
214	CGGGGUUU U UCUUGUUG	70	CAACAAGA CUGAUGAG GCCGUUAGGC CGAA AAACCCCG	7470
215	GGGGUUUU U CUUGUUGA	71	UCAACAAG CUGAUGAG GCCGUUAGGC CGAA AAAACCCC	7471
216	GGGUUUUU C UUGUUGAC	72	GUCAACAA CUGAUGAG GCCGUUAGGC CGAA AAAAACCC	7472
218	GUUUUUCU U GUUGACAA	73	UUGUCAAC CUGAUGAG GCCGUUAGGC CGAA AGAAAAAC	7473
221	UUUCUUGU U GACAAAAA	74	UUUUUGUC CUGAUGAG GCCGUUAGGC CGAA ACAAGAAA	7474
231	ACAAAAAU C CUCACAAU	75	AUUGUGAG CUGAUGAG GCCGUUAGGC CGAA AUUUUUGU	7475
234	AAAAUCCU C ACAAUACC	76	GGUAUUGU CUGAUGAG GCCGUUAGGC CGAA AGGAUUUU	7476
240	CUCACAAU A CCACAGAG	77	CUCUGUGG CUGAUGAG GCCGUUAGGC CGAA AUUGUGAG	7477
250	CACAGAGU C UAGACUCG	78	CGAGUCUA CUGAUGAG GCCGUUAGGC CGAA ACUCUGUG	7478
252	CAGAGUCU A GACUCGUG	79	CACGAGUC CUGAUGAG GCCGUUAGGC CGAA AGACUCUG	7479

257	UCUAGACU C GUGGUGGA	80	UCCACCAC CUGAUGAG GCCGUUAGGC CGAA AGUCUAGA	7480
268	GGUGGACU U CUCUCAAU	81	AUUGAGAG CUGAUGAG GCCGUUAGGC CGAA AGUCCACC	7481
269	GUGGACUU C UCUCAAUU	82	AAUUGAGA CUGAUGAG GCCGUUAGGC CGAA AAGUCCAC	7482
271	GGACUUCU C UCAAUUUU	83	AAAAUUGA CUGAUGAG GCCGUUAGGC CGAA AGAAGUCC	7483
273	ACUUCUCU C AAUUUUCU	84	AGAAAAUU CUGAUGAG GCCGUUAGGC CGAA AGAGAAGU	7484
277	CUCUCAAU U UUCUAGGG	85	CCCUAGAA CUGAUGAG GCCGUUAGGC CGAA AUUGAGAG	7485
278	UCUCAAUU U UCUAGGGG	86	CCCCUAGA CUGAUGAG GCCGUUAGGC CGAA AAUUGAGA	7486
279	CUCAAUUU U CUAGGGGG	87	CCCCCUAG CUGAUGAG GCCGUUAGGC CGAA AAAUUGAG	7487
280	UCAAUUUU C UAGGGGGA	88	UCCCCCUA CUGAUGAG GCCGUUAGGC CGAA AAAAUUGA	7488
282	AAUUUUCU A GGGGGAAC	89	GUUCCCCC CUGAUGAG GCCGUUAGGC CGAA AGAAAAUU	7489
301	CCGUGUGU C UUGGCCAA	90	UUGGCCAA CUGAUGAG GCCGUUAGGC CGAA ACACACGG	7490
303	GUGUGUCU U GGCCAAAA	91	UUUUGGCC CUGAUGAG GCCGUUAGGC CGAA AGACACAC	7491
313	GCCAAAAU U CGCAGUCC	92	GGACUGCG CUGAUGAG GCCGUUAGGC CGAA AUUUUGGC	7492
314	CCAAAAUU C GCAGUCCC	93	GGGACUGC CUGAUGAG GCCGUUAGGC CGAA AAUUUUGG	7493
320	UUCGCAGU C CCAAAUCU	94	AGAUUUGG CUGAUGAG GCCGUUAGGC CGAA ACUGCGAA	7494
327	UCCCAAAU C UCCAGUCA	95	UGACUGGA CUGAUGAG GCCGUUAGGC CGAA AUUUGGGA	7495
329	CCAAAUCU C CAGUCACU	96	AGUGACUG CUGAUGAG GCCGUUAGGC CGAA AGAUUUGG	7496
334	UCUCCAGU C ACUCACCA	97	UGGUGAGU CUGAUGAG GCCGUUAGGC CGAA ACUGGAGA	7497
338	CAGUCACU C ACCAACCU	98	AGGUUGGU CUGAUGAG GCCGUUAGGC CGAA AGUGACUG	7498
349	CAACCUGU U GUCCUCCA	99	UGGAGGAC CUGAUGAG GCCGUUAGGC CGAA ACAGGUUG	7499
352	CCUGUUGU C CUCCAAUU	100	AAUUGGAG CUGAUGAG GCCGUUAGGC CGAA ACAACAGG	7500
355	GUUGUCCU C CAAUUUGU	101	ACAAAUUG CUGAUGAG GCCGUUAGGC CGAA AGGACAAC	7501
360	CCUCCAAU U UGUCCUGG	102	CCAGGACA CUGAUGAG GCCGUUAGGC CGAA AUUGGAGG	7502
361	CUCCAAUU U GUCCUGGU	103	ACCAGGAC CUGAUGAG GCCGUUAGGC CGAA AAUUGGAG	7503
364	CAAUUUGU C CUGGUUAU	104	AUAACCAG CUGAUGAG GCCGUUAGGC CGAA ACAAAUUG	7504
370	GUCCUGGU U AUCGCUGG	105	CCAGCGAU CUGAUGAG GCCGUUAGGC CGAA ACCAGGAC	7505
371	UCCUGGUU A UCGCUGGA	106	UCCAGCGA CUGAUGAG GCCGUUAGGC CGAA AACCAGGA	7506
373	CUGGUUAU C GCUGGAUG	107	CAUCCAGC CUGAUGAG GCCGUUAGGC CGAA AUAACCAG	7507
385	GGAUGUGU C UGCGGCGU	108	ACGCCGCA CUGAUGAG GCCGUUAGGC CGAA ACACAUCC	7508
394	UGCGGCGU U UUAUCAUC	109	GAUGAUAA CUGAUGAG GCCGUUAGGC CGAA ACGCCGCA	7509
395	GCGGCGUU U UAUCAUCU	110	AGAUGAUA CUGAUGAG GCCGUUAGGC CGAA AACGCCGC	7510
396	CGGCGUUU U AUCAUCUU	111	AAGAUGAU CUGAUGAG GCCGUUAGGC CGAA AAACGCCG	7511
397	GGCGUUUU A UCAUCUUC	112	GAAGAUGA CUGAUGAG GCCGUUAGGC CGAA AAAACGCC	7512
399	CGUUUUAU C AUCUUCCU	113	AGGAAGAU CUGAUGAG GCCGUUAGGC CGAA AUAAAACG	7513
402	UUUAUCAU C UUCCUCUG	114	CAGAGGAA CUGAUGAG GCCGUUAGGC CGAA AUGAUAAA	7514
404	UAUCAUCU U CCUCUGCA	115	UGCAGAGG CUGAUGAG GCCGUUAGGC CGAA AGAUGAUA	7515
405	AUCAUCUU C CUCUGCAU	116	AUGCAGAG CUGAUGAG GCCGUUAGGC CGAA AAGAUGAU	7516
408	AUCUUCCU C UGCAUCCU	117	AGGAUGCA CUGAUGAG GCCGUUAGGC CGAA AGGAAGAU	7517
414	CUCUGCAU C CUGCUGCU	118	AGCAGCAG CUGAUGAG GCCGUUAGGC CGAA AUGCAGAG	7518
423	CUGCUGCU A UGCCUCAU	119	AUGAGGCA CUGAUGAG GCCGUUAGGC CGAA AGCAGCAG	7519
429	CUAUGCCU C AUCUUCUU	120	AAGAAGAU CUGAUGAG GCCGUUAGGC CGAA AGGCAUAG	7520
432	UGCCUCAU C UUCUUGUU	121	AACAAGAA CUGAUGAG GCCGUUAGGC CGAA AUGAGGCA	7521
434	CCUCAUCU U CUUGUUGG	122	CCAACAAG CUGAUGAG GCCGUUAGGC CGAA AGAUGAGG	7522
435	CUCAUCUU C UUGUUGGU	123	ACCAACAA CUGAUGAG GCCGUUAGGC CGAA AAGAUGAG	7523
437	CAUCUUCU U GUUGGUUC	124	GAACCAAC CUGAUGAG GCCGUUAGGC CGAA AGAAGAUG	7524
440	CUUCUUGU U GGUUCUUC	125	GAAGAACC CUGAUGAG GCCGUUAGGC CGAA ACAAGAAG	7525
444	UUGUUGGU U CUUCUGGA	126	UCCAGAAG CUGAUGAG GCCGUUAGGC CGAA ACCAACAA	7526
445	UGUUGGUU C UUCUGGAC	127	GUCCAGAA CUGAUGAG GCCGUUAGGC CGAA AACCAACA	7527
447	UUGGUUCU U CUGGACUA	128	UAGUCCAG CUGAUGAG GCCGUUAGGC CGAA AGAACCAA	7528
448	UGGUUCUU C UGGACUAU	129	AUAGUCCA CUGAUGAG GCCGUUAGGC CGAA AAGAACCA	7529
455	UCUGGACU A UCAAGGUA	130	UACCUUGA CUGAUGAG GCCGUUAGGC CGAA AGUCCAGA	7530

457	UGGACUAU C AAGGUAUG	131	CAUACCUU CUGAUGAG GCCGUUAGGC CGAA AUAGUCCA	7531
463	AUCAAGGU A UGUUGCCC	132	GGGCAACA CUGAUGAG GCCGUUAGGC CGAA ACCUUGAU	7532
467	AGGUAUGU U GCCCGUUU	133	AAACGGGC CUGAUGAG GCCGUUAGGC CGAA ACAUACCU	7533
474	UUGCCCGU U UGUCCUCU	134	AGAGGACA CUGAUGAG GCCGUUAGGC CGAA ACGGGCAA	7534
475	UGCCCGUU U GUCCUCUA	135	UAGAGGAC CUGAUGAG GCCGUUAGGC CGAA AACGGGCA	7535
478	CCGUUUGU C CUCUAAUU	136	AAUUAGAG CUGAUGAG GCCGUUAGGC CGAA ACAAACGG	7536
481	UUUGUCCU C UAAUUCCA	137	UGGAAUUA CUGAUGAG GCCGUUAGGC CGAA AGGACAAA	7537
483	UGUCCUCU A AUUCCAGG	138	CCUGGAAU CUGAUGAG GCCGUUAGGC CGAA AGAGGACA	7538
486	CCUCUAAU U CCAGGAUC	139	GAUCCUGG CUGAUGAG GCCGUUAGGC CGAA AUUAGAGG	7539
487	CUCUAAUU C CAGGAUCA	140	UGAUCCUG CUGAUGAG GCCGUUAGGC CGAA AAUUAGAG	7540
494	UCCAGGAU C AUCAACAA	141	UUGUUGAU CUGAUGAG GCCGUUAGGC CGAA AUCCUGGA	7541
497	AGGAUCAU C AACAACCA	142	UGGUUGUU CUGAUGAG GCCGUUAGGC CGAA AUGAUCCU	7542
535	GCACAACU C CUGCUCAA	143	UUGAGCAG CUGAUGAG GCCGUUAGGC CGAA AGUUGUGC	7543
541	CUCCUGCU C AAGGAACC	144	GGUUCCUU CUGAUGAG GCCGUUAGGC CGAA AGCAGGAG	7544
551	AGGAACCU C UAUGUUUC	145	GAAACAUA CUGAUGAG GCCGUUAGGC CGAA AGGUUCCU	7545
553	GAACCUCU A UGUUUCCC	146	GGGAAACA CUGAUGAG GCCGUUAGGC CGAA AGAGGUUC	7546
557	CUCUAUGU U UCCCUCAU	147	AUGAGGGA CUGAUGAG GCCGUUAGGC CGAA ACAUAGAG	7547
558	UCUAUGUU U CCCUCAUG·	148	CAUGAGGG CUGAUGAG GCCGUUAGGC CGAA AACAUAGA	7548
559	CUAUGUUU C CCUCAUGU	149	ACAUGAGG CUGAUGAG GCCGUUAGGC CGAA AAACAUAG	7549
563	GUUUCCCU C AUGUUGCU	150	AGCAACAU CUGAUGAG GCCGUUAGGC CGAA AGGGAAAC	7550
568	CCUCAUGU U GCUGUACA	151	UGUACAGC CUGAUGAG GCCGUUAGGC CGAA ACAUGAGG	7551
574	GUUGCUGU A CAAAACCU	152	AGGUUUUG CUGAUGAG GCCGUUAGGC CGAA ACAGCAAC	7552
583	CAAAACCU A CGGACGGA	153	UCCGUCCG CUGAUGAG GCCGUUAGGC CGAA AGGUUUUG	7553
604	GCACCUGU A UUCCCAUC	154	GAUGGGAA CUGAUGAG GCCGUUAGGC CGAA ACAGGUGC	7554
606	ACCUGUAU U CCCAUCCC	155	GGGAUGGG CUGAUGAG GCCGUUAGGC CGAA AUACAGGU	7555
607	CCUGUAUU C CCAUCCCA	156	UGGGAUGG CUGAUGAG GCCGUUAGGC CGAA AAUACAGG	7556
612	AUUCCCAU C CCAUCAUC	157	GAUGAUGG CUGAUGAG GCCGUUAGGC CGAA AUGGGAAU	7557
617	CAUCCCAU C AUCUUGGG	158	CCCAAGAU CUGAUGAG GCCGUUAGGC CGAA AUGGGAUG	7558
620	CCCAUCAU C UUGGGCUU	159	AAGCCCAA CUGAUGAG GCCGUUAGGC CGAA AUGAUGGG	7559
622	CAUCAUCU U GGGCUUUC	160	GAAAGCCC CUGAUGAG GCCGUUAGGC CGAA AGAUGAUG	7560
628	CUUGGGCU U UCGCAAAA	161	UUUUGCGA CUGAUGAG GCCGUUAGGC CGAA AGCCCAAG	7561
629	UUGGGCUU U CGCAAAAU	162	AUUUUGCG CUGAUGAG GCCGUUAGGC CGAA AAGCCCAA	7562
630	UGGGCUUU C GCAAAAUA	163	UAUUUUGC CUGAUGAG GCCGUUAGGC CGAA AAAGCCCA	7563
638	CGCAAAAU A CCUAUGGG	164	CCCAUAGG CUGAUGAG GCCGUUAGGC CGAA AUUUUGCG	7564
642	AAAUACCU A UGGGAGUG	165	CACUCCCA CUGAUGAG GCCGUUAGGC CGAA AGGUAUUU	7565
656	GUGGGCCU C AGUCCGUU	166	AACGGACU CUGAUGAG GCCGUUAGGC CGAA AGGCCCAC	7566
660	GCCUCAGU C CGUUUCUC	167	GAGAAACG CUGAUGAG GCCGUUAGGC CGAA ACUGAGGC	7567
664	CAGUCCGU U UCUCUUGG	168	CCAAGAGA CUGAUGAG GCCGUUAGGC CGAA ACGGACUG	7568
665	AGUCCGUU U CUCUUGGC	169	GCCAAGAG CUGAUGAG GCCGUUAGGC CGAA AACGGACU	7569
666	GUCCGUUU C UCUUGGCU	170	AGCCAAGA CUGAUGAG GCCGUUAGGC CGAA AAACGGAC	7570
668	CCGUUUCU C UUGGCUCA	171	UGAGCCAA CUGAUGAG GCCGUUAGGC CGAA AGAAACGG	7571
670	GUUUCUCU U GGCUCAGU	172	ACUGAGCC CUGAUGAG GCCGUUAGGC CGAA AGAGAAAC	7572
675	UCUUGGCU C AGUUUACU	173	AGUAAACU CUGAUGAG GCCGUUAGGC CGAA AGCCAAGA	7573
679	GGCUCAGU U UACUAGUG	174	CACUAGUA CUGAUGAG GCCGUUAGGC CGAA ACUGAGCC	7574
680	GCUCAGUU U ACUAGUGC	175	GCACUAGU CUGAUGAG GCCGUUAGGC CGAA AACUGAGC	7575
681	CUCAGUUU A CUAGUGCC	176	GGCACUAG CUGAUGAG GCCGUUAGGC CGAA AAACUGAG	7576
684	AGUUUACU A GUGCCAUU	177	AAUGGCAC CUGAUGAG GCCGUUAGGC CGAA AGUAAACU	7577
692	AGUGCCAU U UGUUCAGU	178	ACUGAACA CUGAUGAG GCCGUUAGGC CGAA AUGGCACU	7578
693	GUGCCAUU U GUUCAGUG	179	CACUGAAC CUGAUGAG GCCGUUAGGC CGAA AAUGGCAC	7579
696	CCAUUUGU U CAGUGGUU	180	AACCACUG CUGAUGAG GCCGUUAGGC CGAA ACAAAUGG	7580
697	CAUUUGUU C AGUGGUUC	181	GAACCACU CUGAUGAG GCCGUUAGGC CGAA AACAAAUG	7581

704	UCAGUGGU U CGUAGGGC	182	GCCCUACG CUGAUGAG GCCGUUAGGC CGAA ACCACUGA	7582
705	CAGUGGUU C GUAGGGCU	183	AGCCCUAC CUGAUGAG GCCGUUAGGC CGAA AACCACUG	7583
708	UGGUUCGU A GGGCUUUC	184	GAAAGCCC CUGAUGAG GCCGUUAGGC CGAA ACGAACCA	7584
714	GUAGGGCU U UCCCCCAC	185	GUGGGGGA CUGAUGAG GCCGUUAGGC CGAA AGCCCUAC	7585
715	UAGGGCUU U CCCCCACU	186	AGUGGGGG CUGAUGAG GCCGUUAGGC CGAA AAGCCCUA	7586
716	AGGGCUUU C CCCCACUG	187	CAGUGGGG CUGAUGAG GCCGUUAGGC CGAA AAAGCCCU	7587
726	CCCACUGU C UGGCUUUC	188	GAAAGCCA CUGAUGAG GCCGUUAGGC CGAA ACAGUGGG	7588
732	GUCUGGCU U UCAGUUAU	189	AUAACUGA CUGAUGAG GCCGUUAGGC CGAA AGCCAGAC	7589
733	UCUGGCUU U CAGUUAUA	190	UAUAACUG CUGAUGAG GCCGUUAGGC CGAA AAGCCAGA	7590
734	CUGGCUUU C AGUUAUAU	191	AUAUAACU CUGAUGAG GCCGUUAGGC CGAA AAAGCCAG	7591
738	CUUUCAGU U AUAUGGAU	192	AUCCAUAU CUGAUGAG GCCGUUAGGC CGAA ACUGAAAG	7592
739	UUUCAGUU A UAUGGAUG	193	CAUCCAUA CUGAUGAG GCCGUUAGGC CGAA AACUGAAA	7593
741	UCAGUUAU A UGGAUGAU	194	AUCAUCCA CUGAUGAG GCCGUUAGGC CGAA AUAACUGA	7594
755	GAUGUGGU U UUGGGGGC	195	GCCCCCAA CUGAUGAG GCCGUUAGGC CGAA ACCACAUC	7595
756	AUGUGGUU U UGGGGGCC	196	GGCCCCCA CUGAUGAG GCCGUUAGGC CGAA AACCACAU	7596
757	UGUGGUUU U GGGGGCCA	197	UGGCCCCC CUGAUGAG GCCGUUAGGC CGAA AAACCACA	7597
769	GGCCAAGU C UGUACAAC	198	GUUGUACA CUGAUGAG GCCGUUAGGC CGAA ACUUGGCC	7598
773	AAGUCUGU A CAACAUCU	199	AGAUGUUG CUGAUGAG GCCGUUAGGC CGAA ACAGACUU	7599
780	UACAACAU C UUGAGUCC	200	GGACUCAA CUGAUGAG GCCGUUAGGC CGAA AUGUUGUA	7600
782	CAACAUCU U GAGUCCCU	201	AGGGACUC CUGAUGAG GCCGUUAGGC CGAA AGAUGUUG	7601
787	UCUUGAGU C CCUUUAUG	202	CAUAAAGG CUGAUGAG GCCGUUAGGC CGAA ACUCAAGA	7602
791	GAGUCCCU U UAUGCCGC	203	GCGGCAUA CUGAUGAG GCCGUUAGGC CGAA AGGGACUC	7603
792	AGUCCCUU U AUGCCGCU	204	AGCGGCAU CUGAUGAG GCCGUUAGGC CGAA AAGGGACU	7604
793	GUCCCUUU A UGCCGCUG	205	CAGCGGCA CUGAUGAG GCCGUUAGGC CGAA AAAGGGAC	7605
803	GCCGCUGU U ACCAAUUU	206	AAAUUGGU CUGAUGAG GCCGUUAGGC CGAA ACAGCGGC	7606
804	CCGCUGUU A CCAAUUUU	207	AAAAUUGG CUGAUGAG GCCGUUAGGC CGAA AACAGCGG	7607
810	UUACCAAU U UUCUUUUG	208	CAAAAGAA CUGAUGAG GCCGUUAGGC CGAA AUUGGUAA	7608
811	UACCAAUU U UCUUUUGU	209	ACAAAAGA CUGAUGAG GCCGUUAGGC CGAA AAUUGGUA	7609
812	ACCAAUUU U CUUUUGUC	210	GACAAAAG CUGAUGAG GCCGUUAGGC CGAA AAAUUGGU	7610
813	CCAAUUUU C UUUUGUCU	211	AGACAAAA CUGAUGAG GCCGUUAGGC CGAA AAAAUUGG	7611
815	AAUUUUCU U UUGUCUUU	212	AAAGACAA CUGAUGAG GCCGUUAGGC CGAA AGAAAAUU	7612
816	AUUUUCUU U UGUCUUUG	213	CAAAGACA CUGAUGAG GCCGUUAGGC CGAA AAGAAAAU	7613
817	UUUUCUUU U GUCUUUGG	214	CCAAAGAC CUGAUGAG GCCGUUAGGC CGAA AAAGAAAA	7614
820	UCUUUUGU C UUUGGGUA	215	UACCCAAA CUGAUGAG GCCGUUAGGC CGAA ACAAAAGA	7615
822	UUUUGUCU U UGGGUAUA	216	UAUACCCA CUGAUGAG GCCGUUAGGC CGAA AGACAAAA	7616
823	UUUGUCUU U GGGUAUAC	217	GUAUACCC CUGAUGAG GCCGUUAGGC CGAA AAGACAAA	7617
828	CUUUGGGU A UACAUUUA	218	UAAAUGUA CUGAUGAG GCCGUUAGGC CGAA ACCCAAAG	7618
830	UUGGGUAU A CAUUUAAA	219	UUUAAAUG CUGAUGAG GCCGUUAGGC CGAA AUACCCAA	7619
834	GUAUACAU U UAAACCCU	220	AGGGUUUA CUGAUGAG GCCGUUAGGC CGAA AUGUAUAC	7620
835	UAUACAUU U AAACCCUC	221	GAGGGUUU CUGAUGAG GCCGUUAGGC CGAA AAUGUAUA	7621
836	AUACAUUU A AACCCUCA	222	UGAGGGUU CUGAUGAG GCCGUUAGGC CGAA AAAUGUAU	7622
843	UAAACCCU C ACAAAACA	223	UGUUUUGU CUGAUGAG GCCGUUAGGC CGAA AGGGUUUA	7623
865	AUGGGGAU A UUCCCUUA	224	UAAGGGAA CUGAUGAG GCCGUUAGGC CGAA AUCCCCAU	7624
867	GGGGAUAU U CCCUUAAC	225	GUUAAGGG CUGAUGAG GCCGUUAGGC CGAA AUAUCCCC	7625
868	GGGAUAUU C CCUUAACU	226	AGUUAAGG CUGAUGAG GCCGUUAGGC CGAA AAUAUCCC	7626
872	UAUUCCCU U AACUUCAU	227	AUGAAGUU CUGAUGAG GCCGUUAGGC CGAA AGGGAAUA	7627
873	AUUCCCUU A ACUUCAUG	228	CAUGAAGU CUGAUGAG GCCGUUAGGC CGAA AAGGGAAU	7628
877	CCUUAACU U CAUGGGAU	229	AUCCCAUG CUGAUGAG GCCGUUAGGC CGAA AGUUAAGG	7629
878	CUUAACUU C AUGGGAUA	230	UAUCCCAU CUGAUGAG GCCGUUAGGC CGAA AAGUUAAG	7630
886	CAUGGGAU A UGUAAUUG	231	CAAUUACA CUGAUGAG GCCGUUAGGC CGAA AUCCCAUG	7631
890	GGAUAUGU A AUUGGGAG	232	CUCCCAAU CUGAUGAG GCCGUUAGGC CGAA ACAUAUCC	7632

893	UAUGUAAU U GGGAGUUG	233	CAACUCCC CUGAUGAG GCCGUUAGGC CGAA AUUACAUA	7633
900	UUGGGAGU U GGGGCACA	234	UGUGCCCC CUGAUGAG GCCGUUAGGC CGAA ACUCCCAA	7634
910	GGGCACAU U GCCACAGG	235	CCUGUGGC CUGAUGAG GCCGUUAGGC CGAA AUGUGCCC	7635
924	AGGAACAU A UUGUACAA	236	UUGUACAA CUGAUGAG GCCGUUAGGC CGAA AUGUUCCU	7636
926	GAACAUAU U GUACAAAA	237	UUUUGUAC CUGAUGAG GCCGUUAGGC CGAA AUAUGUUC	7637
929	CAUAUUGU A CAAAAAAU	238	AUUUUUUG CUGAUGAG GCCGUUAGGC CGAA ACAAUAUG	7638
938	CAAAAAAU C AAAAUGUG	239	CACAUUUU CUGAUGAG GCCGUUAGGC CGAA AUUUUUUG	7639
948	AAAUGUGU U UUAGGAAA	240	UUUCCUAA CUGAUGAG GCCGUUAGGC CGAA ACACAUUU	7640
949	AAUGUGUU U UAGGAAAC	241	GUUUCCUA CUGAUGAG GCCGUUAGGC CGAA AACACAUU	7641
950	AUGUGUUU U AGGAAACU	242	AGUUUCCU CUGAUGAG GCCGUUAGGC CGAA AAACACAU	7642
951	UGUGUUUU A GGAAACUU	243	AAGUUUCC CUGAUGAG GCCGUUAGGC CGAA AAAACACA	7643
959	AGGAAACU U CCUGUAAA	244	UUUACAGG CUGAUGAG GCCGUUAGGC CGAA AGUUUCCU	7644
960	GGAAACUU C CUGUAAAC	245	GUUUACAG CUGAUGAG GCCGUUAGGC CGAA AAGUUUCC	7645
965	CUUCCUGU A AACAGGCC	246	GGCCUGUU CUGAUGAG GCCGUUAGGC CGAA ACAGGAAG	7646
975	ACAGGCCU A UUGAUUGG	247	CCAAUCAA CUGAUGAG GCCGUUAGGC CGAA AGGCCUGU	7647
977	AGGCCUAU U GAUUGGAA	248	UUCCAAUC CUGAUGAG GCCGUUAGGC CGAA AUAGGCCU	7648
981	CUAUUGAU U GGAAAGUA	249	UACUUUCC CUGAUGAG GCCGUUAGGC CGAA AUCAAUAG	7649
989	UGGAAAGU A UGUCAACG	250	CGUUGACA CUGAUGAG GCCGUUAGGC CGAA ACUUUCCA	7650
993	AAGUAUGU C AACGAAUU	251	AAUUCGUU CUGAUGAG GCCGUUAGGC CGAA ACAUACUU	7651
1001	CAACGAAU U GUGGGUCU	252	AGACCCAC CUGAUGAG GCCGUUAGGC CGAA AUUCGUUG	7652
1008	UUGUGGGU C UUUUGGGG	253	CCCCAAAA CUGAUGAG GCCGUUAGGC CGAA ACCCACAA	7653
1010	GUGGGUCU U UUGGGGUU	254	AACCCCAA CUGAUGAG GCCGUUAGGC CGAA AGACCCAC	7654
1011	UGGGUCUU U UGGGGUUU	255	AAACCCCA CUGAUGAG GCCGUUAGGC CGAA AAGACCCA	7655
1012	GGGUCUUU U GGGGUUUG	256	CAAACCCC CUGAUGAG GCCGUUAGGC CGAA AAAGACCC	7656
1018	UUUGGGGU U UGCCGCCC	257	GGGCGGCA CUGAUGAG GCCGUUAGGC CGAA ACCCCAAA	7657
1019	UUGGGGUU U GCCGCCCC	258	GGGGCGGC CUGAUGAG GCCGUUAGGC CGAA AACCCCAA	7658
1029	CCGCCCCU U UCACGCAA	259	UUGCGUGA CUGAUGAG GCCGUUAGGC CGAA AGGGGCGG	7659
1030	CGCCCCUU U CACGCAAU	260	AUUGCGUG CUGAUGAG GCCGUUAGGC CGAA AAGGGGCG	7660
1031	GCCCCUUU C ACGCAAUG	261	CAUUGCGU CUGAUGAG GCCGUUAGGC CGAA AAAGGGGC	7661
1045	AUGUGGAU A UUCUGCUU	262	AAGCAGAA CUGAUGAG GCCGUUAGGC CGAA AUCCACAU	7662
1047	GUGGAUAU U CUGCUUUA	263	UAAAGCAG CUGAUGAG GCCGUUAGGC CGAA AUAUCCAC	7663
1048	UGGAUAUU C UGCUUUAA	264	UUAAAGCA CUGAUGAG GCCGUUAGGC CGAA AAUAUCCA	7664
1053	AUUCUGCU U UAAUGCCU	265	AGGCAUUA CUGAUGAG GCCGUUAGGC CGAA AGCAGAAU	7665
1054	UUCUGCUU U AAUGCCUU	266	AAGGCAUU CUGAUGAG GCCGUUAGGC CGAA AAGCAGAA	7666
1055	UCUGCUUU A AUGCCUUU	267	AAAGGCAU CUGAUGAG GCCGUUAGGC CGAA AAAGCAGA	7667
1062	UAAUGCCU U UAUAUGCA	268	UGCAUAUA CUGAUGAG GCCGUUAGGC CGAA AGGCAUUA	7668
1063	AAUGCCUU U AUAUGCAU	269	AUGCAUAU CUGAUGAG GCCGUUAGGC CGAA AAGGCAUU	7669
1064	AUGCCUUU A UAUGCAUG	270	CAUGCAUA CUGAUGAG GCCGUUAGGC CGAA AAAGGCAU	7670
1066	GCCUUUAU A UGCAUGCA	271	UGCAUGCA CUGAUGAG GCCGUUAGGC CGAA AUAAAGGC	7671
1076	GCAUGCAU A CAAGCAAA	272	UUUGCUUG CUGAUGAG GCCGUUAGGC CGAA AUGCAUGC	7672
1092	AACAGGCU U UUACUUUC	273	GAAAGUAA CUGAUGAG GCCGUUAGGC CGAA AGCCUGUU	7673
1093	ACAGGCUU U UACUUUCU	274	AGAAAGUA CUGAUGAG GCCGUUAGGC CGAA AAGCCUGU	7674
1094	CAGGCUUU U ACUUUCUC	275	GAGAAAGU CUGAUGAG GCCGUUAGGC CGAA AAAGCCUG	7675
1095	AGGCUUUU A CUUUCUCG	276	CGAGAAAG CUGAUGAG GCCGUUAGGC CGAA AAAAGCCU	7676
1098	CUUUUACU U UCUCGCCA	277	UGGCGAGA CUGAUGAG GCCGUUAGGC CGAA AGUAAAAG	7677
1099	UUUUACUU U CUCGCCAA	278	UUGGCGAG CUGAUGAG GCCGUUAGGC CGAA AAGUAAAA	7678
1100	UUUACUUU C UCGCCAAC	279	GUUGGCGA CUGAUGAG GCCGUUAGGC CGAA AAAGUAAA	7679
1102	UACUUUCU C GCCAACUU	280	AAGUUGGC CUGAUGAG GCCGUUAGGC CGAA AGAAAGUA	7680
1110	CGCCAACU U ACAAGGCC	281	GGCCUUGU CUGAUGAG GCCGUUAGGC CGAA AGUUGGCG	7681
1111	GCCAACUU A CAAGGCCU	282	AGGCCUUG CUGAUGAG GCCGUUAGGC CGAA AAGUUGGC	7682
1120	CAAGGCCU U UCUAAGUA	283	UACUUAGA CUGAUGAG GCCGUUAGGC CGAA AGGCCUUG	7683

1121	AAGGCCUU U CUAAGUAA	284	UUACUUAG CUGAUGAG GCCGUUAGGC CGAA AAGGCCUU	7684
1122	AGGCCUUU C UAAGUAAA	285	UUUACUUA CUGAUGAG GCCGUUAGGC CGAA AAAGGCCU	7685
1124	GCCUUUCU A AGUAAACA	286	UGUUUACU CUGAUGAG GCCGUUAGGC CGAA AGAAAGGC	7686
1128	UUCUAAGU A AACAGUAU	287	AUACUGUU CUGAUGAG GCCGUUAGGC CGAA ACUUAGAA	7687
1135	UAAACAGU A UGUGAACC	288	GGUUCACA CUGAUGAG GCCGUUAGGC CGAA ACUGUUUA	7688
1145	GUGAACCU U UACCCCGU	289	ACGGGGUA CUGAUGAG GCCGUUAGGC CGAA AGGUUCAC	7689
1146	UGAACCUU U ACCCCGUU	290	AACGGGGU CUGAUGAG GCCGUUAGGC CGAA AAGGUUCA	7690
1147	GAACCUUU A CCCCGUUG	291	CAACGGGG CUGAUGAG GCCGUUAGGC CGAA AAAGGUUC	7691
1154	UACCCCGU U GCUCGGCA	292	UGCCGAGC CUGAUGAG GCCGUUAGGC CGAA ACGGGGUA	7692
1158	CCGUUGCU C GGCAACGG	293	CCGUUGCC CUGAUGAG GCCGUUAGGC CGAA AGCAACGG	7693
1173	GGCCUGGU C UAUGCCAA	294	UUGGCAUA CUGAUGAG GCCGUUAGGC CGAA ACCAGGCC	7694
1175	CCUGGUCU A UGCCAAGU	295	ACUUGGCA CUGAUGAG GCCGUUAGGC CGAA AGACCAGG	7695
1186	CCAAGUGU U UGCUGACG	296	CGUCAGCA CUGAUGAG GCCGUUAGGC CGAA ACACUUGG	7696
1187	CAAGUGUU U GCUGACGC	297	GCGUCAGC CUGAUGAG GCCGUUAGGC CGAA AACACUUG	7697
1209	CCACUGGU U GGGGCUUG	298	CAAGCCCC CUGAUGAG GCCGUUAGGC CGAA ACCAGUGG	7698
1216	UUGGGGCU U GGCCAUAG	299	CUAUGGCC CUGAUGAG GCCGUUAGGC CGAA AGCCCCAA	7699
1223	UUGGCCAU A GGCCAUCA	300	UGAUGGCC CUGAUGAG GCCGUUAGGC CGAA AUGGCCAA	7700
1230	UAGGCCAU C AGCGCAUG	301	CAUGCGCU CUGAUGAG GCCGUUAGGC CGAA AUGGCCUA	7701
1249	UGGAACCU U UGUGUCUC	302	GAGACACA CUGAUGAG GCCGUUAGGC CGAA AGGUUCCA	7702
1250	GGAACCUU U GUGUCUCC	303	GGAGACAC CUGAUGAG GCCGUUAGGC CGAA AAGGUUCC	7703
1255	CUUUGUGU C UCCUCUGC	304	GCAGAGGA CUGAUGAG GCCGUUAGGC CGAA ACACAAAG	7704
1257	UUGUGUCU C CUCUGCCG	305	CGGCAGAG CUGAUGAG GCCGUUAGGC CGAA AGACACAA	7705
1260	UGUCUCCU C UGCCGAUC	306	GAUCGGCA CUGAUGAG GCCGUUAGGC CGAA AGGAGACA	7706
1268	CUGCCGAU C CAUACCGC	307	GCGGUAUG CUGAUGAG GCCGUUAGGC CGAA AUCGGCAG	7707
1272	CGAUCCAU A CCGCGGAA	308	UUCCGCGG CUGAUGAG GCCGUUAGGC CGAA AUGGAUCG	7708
1283	GCGGAACU C CUAGCCGC	309	GCGGCUAG CUGAUGAG GCCGUUAGGC CGAA AGUUCCGC	7709
1286	GAACUCCU A GCCGCUUG	310	CAAGCGGC CUGAUGAG GCCGUUAGGC CGAA AGGAGUUC	7710
1293	UAGCCGCU U GUUUUGCU	311	AGCAAAAC CUGAUGAG GCCGUUAGGC CGAA AGCGGCUA	7711
1296	CCGCUUGU U UUGCUCGC	312	GCGAGCAA CUGAUGAG GCCGUUAGGC CGAA ACAAGCGG	7712
1297	CGCUUGUU U UGCUCGCA	313	UGCGAGCA CUGAUGAG GCCGUUAGGC CGAA AACAAGCG	7713
1298	GCUUGUUU U GCUCGCAG	314	CUGCGAGC CUGAUGAG GCCGUUAGGC CGAA AAACAAGC	7714
1302	GUUUUGCU C GCAGCAGG	315	CCUGCUGC CUGAUGAG GCCGUUAGGC CGAA AGCAAAAC	7715
1312	CAGCAGGU C UGGGGCAA	316	UUGCCCCA CUGAUGAG GCCGUUAGGC CGAA ACCUGCUG	7716
1325	GCAAAACU C AUCGGGAC	317	GUCCCGAU CUGAUGAG GCCGUUAGGC CGAA AGUUUUGC	7717
1328	AAACUCAU C GGGACUGA	318	UCAGUCCC CUGAUGAG GCCGUUAGGC CGAA AUGAGUUU	7718
1341	CUGACAAU U CUGUCGUG	319	CACGACAG CUGAUGAG GCCGUUAGGC CGAA AUUGUCAG	7719
1342	UGACAAUU C UGUCGUGC	320	GCACGACA CUGAUGAG GCCGUUAGGC CGAA AAUUGUCA	7720
1346	AAUUCUGU C GUGCUCUC	321	GAGAGCAC CUGAUGAG GCCGUUAGGC CGAA ACAGAAUU	7721
1352	GUCGUGCU C UCCCGCAA	322	UUGCGGGA CUGAUGAG GCCGUUAGGC CGAA AGCACGAC	7722
1354	CGUGCUCU C CCGCAAAU	323	AUUUGCGG CUGAUGAG GCCGUUAGGC CGAA AGAGCACG	7723
1363	CCGCAAAU A UACAUCAU	324	AUGAUGUA CUGAUGAG GCCGUUAGGC CGAA AUUUGCGG	7724
1365	GCAAAUAU A CAUCAUUU	325	AAAUGAUG CUGAUGAG GCCGUUAGGC CGAA AUAUUUGC	7725
1369	AUAUACAU C AUUUCCAU	326	AUGGAAAU CUGAUGAG GCCGUUAGGC CGAA AUGUAUAU	7726
1372	UACAUCAU U UCCAUGGC	327	GCCAUGGA CUGAUGAG GCCGUUAGGC CGAA AUGAUGUA	7727
1373	ACAUCAUU U CCAUGGCU	328	AGCCAUGG CUGAUGAG GCCGUUAGGC CGAA AAUGAUGU	7728
1374	CAUCAUUU C CAUGGCUG	329	CAGCCAUG CUGAUGAG GCCGUUAGGC CGAA AAAUGAUG	7729
1385	UGGCUGCU A GGCUGUGC	330	GCACAGCC CUGAUGAG GCCGUUAGGC CGAA AGCAGCCA	7730
1406	AACUGGAU C CUACGCGG	331	CCGCGUAG CUGAUGAG GCCGUUAGGC CGAA AUCCAGUU	7731
1409	UGGAUCCU A CGCGGGAC	332	GUCCCGCG CUGAUGAG GCCGUUAGGC CGAA AGGAUCCA	7732
1420	CGGGACGU C CUUUGUUU	333	AAACAAAG CUGAUGAG GCCGUUAGGC CGAA ACGUCCCG	7733
1423	GACGUCCU U UGUUUACG	334	CGUAAACA CUGAUGAG GCCGUUAGGC CGAA AGGACGUC	7734

1424	ACGUCCUU U GUUUACGU	335	ACGUAAAC CUGAUGAG GCCGUUAGGC CGAA AAGGACGU	7735
1427	UCCUUUGU U UACGUCCC	336	GGGACGUA CUGAUGAG GCCGUUAGGC CGAA ACAAAGGA	7736
1428	CCUUUGUU U ACGUCCCG	337	CGGGACGU CUGAUGAG GCCGUUAGGC CGAA AACAAAGG	7737
1429	CUUUGUUU A CGUCCCGU	338	ACGGGACG CUGAUGAG GCCGUUAGGC CGAA AAACAAAG	7738
1433	GUUUACGU C CCGUCGGC	339	GCCGACGG CUGAUGAG GCCGUUAGGC CGAA ACGUAAAC	7739
1438	CGUCCCGU C GGCGCUGA	340	UCAGCGCC CUGAUGAG GCCGUUAGGC CGAA ACGGGACG	7740
1449	CGCUGAAU C CCGCGGAC	341	GUCCGCGG CUGAUGAG GCCGUUAGGC CGAA AUUCAGCG	7741
1465	CGACCCCU C CCGGGGCC	342	GGCCCCGG CUGAUGAG GCCGUUAGGC CGAA AGGGGUCG	7742
1477	GGGCCGCU U GGGGCUCU	343	AGAGCCCC CUGAUGAG GCCGUUAGGC CGAA AGCGGCCC	7743
1484	UUGGGGCU C UACCGCCC	344	GGGCGGUA CUGAUGAG GCCGUUAGGC CGAA AGCCCCAA	7744
1486	GGGGCUCU A CCGCCCGC	345	GCGGGCGG CUGAUGAG GCCGUUAGGC CGAA AGAGCCCC	7745
1496	CGCCCGCU U CUCCGCCU	346	AGGCGGAG CUGAUGAG GCCGUUAGGC CGAA AGCGGGCG	7746
1497	GCCCGCUU C UCCGCCUA	347	UAGGCGGA CUGAUGAG GCCGUUAGGC CGAA AAGCGGGC	7747
1499	CCGCUUCU C CGCCUAUU	348	AAUAGGCG CUGAUGAG GCCGUUAGGC CGAA AGAAGCGG	7748
1505	CUCCGCCU A UUGUACCG	349	CGGUACAA CUGAUGAG GCCGUUAGGC CGAA AGGCGGAG	7749
1507	CCGCCUAU U GUACCGAC	350	GUCGGUAC CUGAUGAG GCCGUUAGGC CGAA AUAGGCGG	7750
1510	CCUAUUGU A CCGACCGU	351	ACGGUCGG CUGAUGAG GCCGUUAGGC CGAA ACAAUAGG	7751
1519	CCGACCGU C CACGGGGC	352	GCCCCGUG CUGAUGAG GCCGUUAGGC CGAA ACGGUCGG	7752
1534	GCGCACCU C UCUUUACG	353	CGUAAAGA CUGAUGAG GCCGUUAGGC CGAA AGGUGCGC	7753
1536	GCACCUCU C UUUACGCG	354	CGCGUAAA CUGAUGAG GCCGUUAGGC CGAA AGAGGUGC	7754
1538	ACCUCUCU U UACGCGGA	355	UCCGCGUA CUGAUGAG GCCGUUAGGC CGAA AGAGAGGU	7755
1539	CCUCUCUU U ACGCGGAC	356	GUCCGCGU CUGAUGAG GCCGUUAGGC CGAA AAGAGAGG	7756
1540	CUCUCUUU A CGCGGACU	357	AGUCCGCG CUGAUGAG GCCGUUAGGC CGAA AAAGAGAG	7757
1549	CGCGGACU C CCCGUCUG	358	CAGACGGG CUGAUGAG GCCGUUAGGC CGAA AGUCCGCG	7758
1555	CUCCCCGU C UGUGCCUU	359	AAGGCACA CUGAUGAG GCCGUUAGGC CGAA ACGGGGAG	7759
1563	CUGUGCCU U CUCAUCUG	360	CAGAUGAG CUGAUGAG GCCGUUAGGC CGAA AGGCACAG	7760
1564	UGUGCCUU C UCAUCUGC	361	GCAGAUGA CUGAUGAG GCCGUUAGGC CGAA AAGGCACA	7761
1566	UGCCUUCU C AUCUGCCG	362	CGGCAGAU CUGAUGAG GCCGUUAGGC CGAA AGAAGGCA	7762
1569	CUUCUCAU C UGCCGGAC	363	GUCCGGCA CUGAUGAG GCCGUUAGGC CGAA AUGAGAAG	7763
1588	UGUGCACU U CGCUUCAC	364	GUGAAGCG CUGAUGAG GCCGUUAGGC CGAA AGUGCACA	7764
1589	GUGCACUU C GCUUCACC	365	GGUGAAGC CUGAUGAG GCCGUUAGGC CGAA AAGUGCAC	7765
1593	ACUUCGCU U CACCUCUG	366	CAGAGGUG CUGAUGAG GCCGUUAGGC CGAA AGCGAAGU	7766
1594	CUUCGCUU C ACCUCUGC	367	GCAGAGGU CUGAUGAG GCCGUUAGGC CGAA AAGCGAAG	7767
1599	CUUCACCU C UGCACGUC	368	GACGUGCA CUGAUGAG GCCGUUAGGC CGAA AGGUGAAG	7768
1607	CUGCACGU C GCAUGGAG	369	CUCCAUGC CUGAUGAG GCCGUUAGGC CGAA ACGUGCAG	7769
1651	CCCAAGGU C UUGCAUAA	370	UUAUGCAA CUGAUGAG GCCGUUAGGC CGAA ACCUUGGG	7770
1653	CAAGGUCU U GCAUAAGA	371	UCUUAUGC CUGAUGAG GCCGUUAGGC CGAA AGACCUUG	7771
1658	UCUUGCAU A AGAGGACU	372	AGUCCUCU CUGAUGAG GCCGUUAGGC CGAA AUGCAAGA	7772
1667	AGAGGACU C UUGGACUU	373	AAGUCCAA CUGAUGAG GCCGUUAGGC CGAA AGUCCUCU	7773
1669	AGGACUCU U GGACUUUC	374	GAAAGUCC CUGAUGAG GCCGUUAGGC CGAA AGAGUCCU	7774
1675	CUUGGACU U UCAGCAAU	375	AUUGCUGA CUGAUGAG GCCGUUAGGC CGAA AGUCCAAG	7775
1676	UUGGACUU U CAGCAAUG	376	CAUUGCUG CUGAUGAG GCCGUUAGGC CGAA AAGUCCAA	7776
1677	UGGACUUU C AGCAAUGU	377	ACAUUGCU CUGAUGAG GCCGUUAGGC CGAA AAAGUCCA	7777
1686	AGCAAUGU C AACGACCG	378	CGGUCGUU CUGAUGAG GCCGUUAGGC CGAA ACAUUGCU	7778
1699	ACCGACCU U GAGGCAUA	379	UAUGCCUC CUGAUGAG GCCGUUAGGC CGAA AGGUCGGU	7779
1707	UGAGGCAU A CUUCAAAG	380	CUUUGAAG CUGAUGAG GCCGUUAGGC CGAA AUGCCUCA	7780
1710	GGCAUACU U CAAAGACU	381	AGUCUUUG CUGAUGAG GCCGUUAGGC CGAA AGUAUGCC	7781
1711	GCAUACUU C AAAGACUG	382	CAGUCUUU CUGAUGAG GCCGUUAGGC CGAA AAGUAUGC	7782
1725	CUGUGUGU U UAAUGAGU	383	ACUCAUUA CUGAUGAG GCCGUUAGGC CGAA ACACACAG	7783
1726	UGUGUGUU U AAUGAGUG	384	CACUCAUU CUGAUGAG GCCGUUAGGC CGAA AACACACA	7784
1727	GUGUGUUU A AUGAGUGG	385	CCACUCAU CUGAUGAG GCCGUUAGGC CGAA AAACACAC	7785

1743	GGAGGAGU U GGGGGAGG	386	CCUCCCCC CUGAUGAG GCCGUUAGGC CGAA ACUCCUCC	7786
1756	GAGGAGGU U AGGUUAAA	387	UUUAACCU CUGAUGAG GCCGUUAGGC CGAA ACCUCCUC	7787
1757	AGGAGGUU A GGUUAAAG	388	CUUUAACC CUGAUGAG GCCGUUAGGC CGAA AACCUCCU	7788
1761	GGUUAGGU U AAAGGUCU	389	AGACCUUU CUGAUGAG GCCGUUAGGC CGAA ACCUAACC	7789
1762	GUUAGGUU A AAGGUCUU	390	AAGACCUU CUGAUGAG GCCGUUAGGC CGAA AACCUAAC	7790
1768	UUAAAGGU C UUUGUACU	391	AGUACAAA CUGAUGAG GCCGUUAGGC CGAA ACCUUUAA	7791
1770	AAAGGUCU U UGUACUAG	392	CUAGUACA CUGAUGAG GCCGUUAGGC CGAA AGACCUUU	7792
1771	AAGGUCUU U GUACUAGG	393	CCUAGUAC CUGAUGAG GCCGUUAGGC CGAA AAGACCUU	7793
1774	GUCUUUGU A CUAGGAGG	394	CCUCCUAG CUGAUGAG GCCGUUAGGC CGAA ACAAAGAC	7794
1777	UUUGUACU A GGAGGCUG	395	CAGCCUCC CUGAUGAG GCCGUUAGGC CGAA AGUACAAA	7795
1787	GAGGCUGU A GGCAUAAA	396	UUUAUGCC CUGAUGAG GCCGUUAGGC CGAA ACAGCCUC	7796
1793	GUAGGCAU A AAUUGGUG	397	CACCAAUU CUGAUGAG GCCGUUAGGC CGAA AUGCCUAC	7797
1797	GCAUAAAU U GGUGUGUU	398	AACACACC CUGAUGAG GCCGUUAGGC CGAA AUUUAUGC	7798
1805	UGGUGUGU U CACCAGCA	399	UGCUGGUG CUGAUGAG GCCGUUAGGC CGAA ACACACCA	7799
1806	GGUGUGUU C ACCAGCAC	400	GUGCUGGU CUGAUGAG GCCGUUAGGC CGAA AACACACC	7800
1824	AUGCAACU U UUUCACCU	401	AGGUGAAA CUGAUGAG GCCGUUAGGC CGAA AGUUGCAU	7801
1825	UGCAACUU U UUCACCUC	402	GAGGUGAA CUGAUGAG GCCGUUAGGC CGAA AAGUUGCA	7802
1826	GCAACUUU U UCACCUCU	403	AGAGGUGA CUGAUGAG GCCGUUAGGC CGAA AAAGUUGC	7803
1827	CAACUUUU U CACCUCUG	404	CAGAGGUG CUGAUGAG GCCGUUAGGC CGAA AAAAGUUG	7804
1828	AACUUUUU C ACCUCUGC	405	GCAGAGGU CUGAUGAG GCCGUUAGGC CGAA AAAAAGUU	7805
1833	UUUCACCU C UGCCUAAU	406	AUUAGGCA CUGAUGAG GCCGUUAGGC CGAA AGGUGAAA	7806
1839	CUCUGCCU A AUCAUCUC	407	GAGAUGAU CUGAUGAG GCCGUUAGGC CGAA AGGCAGAG	7807
1842	UGCCUAAU C AUCUCAUG	408	CAUGAGAU CUGAUGAG GCCGUUAGGC CGAA AUUAGGCA	7808
1845	CUAAUCAU C UCAUGUUC	409	GAACAUGA CUGAUGAG GCCGUUAGGC CGAA AUGAUUAG	7809
1847	AAUCAUCU C AUGUUCAU	410	AUGAACAU CUGAUGAG GCCGUUAGGC CGAA AGAUGAUU	7810
1852	UCUCAUGU U CAUGUCCU	411	AGGACAUG CUGAUGAG GCCGUUAGGC CGAA ACAUGAGA	7811
1853	CUCAUGUU C AUGUCCUA	412	UAGGACAU CUGAUGAG GCCGUUAGGC CGAA AACAUGAG	7812
1858	GUUCAUGU C CUACUGUU	413	AACAGUAG CUGAUGAG GCCGUUAGGC CGAA ACAUGAAC	7813
1861	CAUGUCCU A CUGUUCAA	414	UUGAACAG CUGAUGAG GCCGUUAGGC CGAA AGGACAUG	7814
1866	CCUACUGU U CAAGCCUC	415	GAGGCUUG CUGAUGAG GCCGUUAGGC CGAA ACAGUAGG	7815
1867	CUACUGUU C AAGCCUCC	416	GGAGGCUU CUGAUGAG GCCGUUAGGC CGAA AACAGUAG	7816
1874	UCAAGCCU C CAAGCUGU	417	ACAGCUUG CUGAUGAG GCCGUUAGGC CGAA AGGCUUGA	7817
1887	CUGUGCCU U GGGUGGCU	418	AGCCACCC CUGAUGAG GCCGUUAGGC CGAA AGGCACAG	7818
1896	GGGUGGCU U UGGGGCAU	419	AUGCCCCA CUGAUGAG GCCGUUAGGC CGAA AGCCACCC	7819
1897	GGUGGCUU U GGGGCAUG	420	CAUGCCCC CUGAUGAG GCCGUUAGGC CGAA AAGCCACC	7820
1911	AUGGACAU U GACCCGUA	421	UACGGGUC CUGAUGAG GCCGUUAGGC CGAA AUGUCCAU	7821
1919	UGACCCGU A UAAAGAAU	422	AUUCUUUA CUGAUGAG GCCGUUAGGC CGAA ACGGGUCA	7822
1921	ACCCGUAU A AAGAAUUU	423	AAAUUCUU CUGAUGAG GCCGUUAGGC CGAA AUACGGGU	7823
1928	UAAAGAAU U UGGAGCUU	424	AAGCUCCA CUGAUGAG GCCGUUAGGC CGAA AUUCUUUA	7824
1929	AAAGAAUU U GGAGCUUC	425	GAAGCUCC CUGAUGAG GCCGUUAGGC CGAA AAUUCUUU	7825
1936	UUGGAGCU U CUGUGGAG	426	CUCCACAG CUGAUGAG GCCGUUAGGC CGAA AGCUCCAA	7826
1937	UGGAGCUU C UGUGGAGU	427	ACUCCACA CUGAUGAG GCCGUUAGGC CGAA AAGCUCCA	7827
1946	UGUGGAGU U ACUCUCUU	428	AAGAGAGU CUGAUGAG GCCGUUAGGC CGAA ACUCCACA	7828
1947	GUGGAGUU A CUCUCUUU	429	AAAGAGAG CUGAUGAG GCCGUUAGGC CGAA AACUCCAC	7829
1950	GAGUUACU C UCUUUUUU	430	AAAAAAGA CUGAUGAG GCCGUUAGGC CGAA AGUAACUC	7830
1952	GUUACUCU C UUUUUUGC	431	GCAAAAAA CUGAUGAG GCCGUUAGGC CGAA AGAGUAAC	7831
1954	UACUCUCU U UUUUGCCU	432	AGGCAAAA CUGAUGAG GCCGUUAGGC CGAA AGAGAGUA	7832
1955	ACUCUCUU U UUUGCCUU	433	AAGGCAAA CUGAUGAG GCCGUUAGGC CGAA AAGAGAGU	7833
1956	CUCUCUUU U UUGCCUUC	434	GAAGGCAA CUGAUGAG GCCGUUAGGC CGAA AAAGAGAG	7834
1957	UCUCUUUU U UGCCUUCU	435	AGAAGGCA CUGAUGAG GCCGUUAGGC CGAA AAAAGAGA	7835
1958	CUCUUUUU U GCCUUCUG	436	CAGAAGGC CUGAUGAG GCCGUUAGGC CGAA AAAAAGAG	7836

1963	UUUUGCCU U CUGACUUC	437	GAAGUCAG CUGAUGAG GCCGUUAGGC CGAA AGGCAAAA	7837
1964	UUUGCCUU C UGACUUCU	438	AGAAGUCA CUGAUGAG GCCGUUAGGC CGAA AAGGCAAA	7838
1970	UUCUGACU U CUUUCCUU	439	AAGGAAAG CUGAUGAG GCCGUUAGGC CGAA AGUCAGAA	7839
1971	UCUGACUU C UUUCCUUC	440	GAAGGAAA CUGAUGAG GCCGUUAGGC CGAA AAGUCAGA	7840
1973	UGACUUCU U UCCUUCUA	441	UAGAAGGA CUGAUGAG GCCGUUAGGC CGAA AGAAGUCA	7841
1974	GACUUCUU U CCUUCUAU	442	AUAGAAGG CUGAUGAG GCCGUUAGGC CGAA AAGAAGUC	7842
1975	ACUUCUUU C CUUCUAUU	443	AAUAGAAG CUGAUGAG GCCGUUAGGC CGAA AAAGAAGU	7843
1978	UCUUUCCU U CUAUUCGA	444	UCGAAUAG CUGAUGAG GCCGUUAGGC CGAA AGGAAAGA	7844
1979	CUUUCCUU C UAUUCGAG	445	CUCGAAUA CUGAUGAG GCCGUUAGGC CGAA AAGGAAAG	7845
1981	UUCCUUCU A UUCGAGAU	446	AUCUCGAA CUGAUGAG GCCGUUAGGC CGAA AGAAGGAA	7846
1983	CCUUCUAU U CGAGAUCU	447	AGAUCUCG CUGAUGAG GCCGUUAGGC CGAA AUAGAAGG	7847
1984	CUUCUAUU C GAGAUCUC	448	GAGAUCUC CUGAUGAG GCCGUUAGGC CGAA AAUAGAAG	7848
1990	UUCGAGAU C UCCUCGAC	449	GUCGAGGA CUGAUGAG GCCGUUAGGC CGAA AUCUCGAA	7849
1992	CGAGAUCU C CUCGACAC	450	GUGUCGAG CUGAUGAG GCCGUUAGGC CGAA AGAUCUCG	7850
1995	GAUCUCCU C GACACCGC	451	GCGGUGUC CUGAUGAG GCCGUUAGGC CGAA AGGAGAUC	7851
2006	CACCGCCU C UGCUCUGU	452	ACAGAGCA CUGAUGAG GCCGUUAGGC CGAA AGGCGGUG	7852
2011	CCUCUGCU C UGUAUCGG	453	CCGAUACA CUGAUGAG GCCGUUAGGC CGAA AGCAGAGG	7853
2015	UGCUCUGU A UCGGGGGG	454	CCCCCCGA CUGAUGAG GCCGUUAGGC CGAA ACAGAGCA	7854
2017	CUCUGUAU C GGGGGGCC	455	GGCCCCCC CUGAUGAG GCCGUUAGGC CGAA AUACAGAG	7855
2027	GGGGGCCU U AGAGUCUC	456	GAGACUCU CUGAUGAG GCCGUUAGGC CGAA AGGCCCCC	7856
2028	GGGGCCUU A GAGUCUCC	457	GGAGACUC CUGAUGAG GCCGUUAGGC CGAA AAGGCCCC	7857
2033	CUUAGAGU C UCCGGAAC	458	GUUCCGGA CUGAUGAG GCCGUUAGGC CGAA ACUCUAAG	7858
2035	UAGAGUCU C CGGAACAU	459	AUGUUCCG CUGAUGAG GCCGUUAGGC CGAA AGACUCUA	7859
2044	CGGAACAU U GUUCACCU	460	AGGUGAAC CUGAUGAG GCCGUUAGGC CGAA AUGUUCCG	7860
2047	AACAUUGU U CACCUCAC	461	GUGAGGUG CUGAUGAG GCCGUUAGGC CGAA ACAAUGUU	7861
2048	ACAUUGUU C ACCUCACC	462	GGUGAGGU CUGAUGAG GCCGUUAGGC CGAA AACAAUGU	7862
2053	GUUCACCU C ACCAUACG	463	CGUAUGGU CUGAUGAG GCCGUUAGGC CGAA AGGUGAAC	7863
2059	CUCACCAU A CGGCACUC	464	GAGUGCCG CUGAUGAG GCCGUUAGGC CGAA AUGGUGAG	7864
2067	ACGGCACU C AGGCAAGC	465	GCUUGCCU CUGAUGAG GCCGUUAGGC CGAA AGUGCCGU	7865
2077	GGCAAGCU A UUCUGUGU	466	ACACAGAA CUGAUGAG GCCGUUAGGC CGAA AGCUUGCC	7866
2079	CAAGCUAU U CUGUGUUG	467	CAACACAG CUGAUGAG GCCGUUAGGC CGAA AUAGCUUG	7867
2080	AAGCUAUU C UGUGUUGG	468	CCAACACA CUGAUGAG GCCGUUAGGC CGAA AAUAGCUU	7868
2086	UUCUGUGU U GGGGUGAG	469	CUCACCCC CUGAUGAG GCCGUUAGGC CGAA ACACAGAA	7869
2096	GGGUGAGU U GAUGAAUC	470	GAUUCAUC CUGAUGAG GCCGUUAGGC CGAA ACUCACCC	7870
2104	UGAUGAAU C UAGCCACC	471	GGUGGCUA CUGAUGAG GCCGUUAGGC CGAA AUUCAUCA	7871
2106	AUGAAUCU A GCCACCUG	472	CAGGUGGC CUGAUGAG GCCGUUAGGC CGAA AGAUUCAU	7872
2125	UGGGAAGU A AUUUGGAA	473	UUCCAAAU CUGAUGAG GCCGUUAGGC CGAA ACUUCCCA	7873
2128	GAAGUAAU U UGGAAGAU	474	AUCUUCCA CUGAUGAG GCCGUUAGGC CGAA AUUACUUC	7874
2129	AAGUAAUU U GGAAGAUC	475	GAUCUUCC CUGAUGAG GCCGUUAGGC CGAA AAUUACUU	7875
2137	UGGAAGAU C CAGCAUCC	476	GGAUGCUG CUGAUGAG GCCGUUAGGC CGAA AUCUUCCA	7876
2144	UCCAGCAU C CAGGGAAU	477	AUUCCCUG CUGAUGAG GCCGUUAGGC CGAA AUGCUGGA	7877
2153	CAGGGAAU U AGUAGUCA	478	UGACUACU CUGAUGAG GCCGUUAGGC CGAA AUUCCCUG	7878
2154	AGGGAAUU A GUAGUCAG	479	CUGACUAC CUGAUGAG GCCGUUAGGC CGAA AAUUCCCU	7879
2157	GAAUUAGU A GUCAGCUA	480	UAGCUGAC CUGAUGAG GCCGUUAGGC CGAA ACUAAUUC	7880
2160	UUAGUAGU C AGCUAUGU	481	ACAUAGCU CUGAUGAG GCCGUUAGGC CGAA ACUACUAA	7881
2165	AGUCAGCU A UGUCAACG	482	CGUUGACA CUGAUGAG GCCGUUAGGC CGAA AGCUGACU	7882
2169	AGCUAUGU C AACGUUAA	483	UUAACGUU CUGAUGAG GCCGUUAGGC CGAA ACAUAGCU	7883
2175	GUCAACGU U AAUAUGGG	484	CCCAUAUU CUGAUGAG GCCGUUAGGC CGAA ACGUUGAC	7884
2176	UCAACGUU A AUAUGGGC	485	GCCCAUAU CUGAUGAG GCCGUUAGGC CGAA AACGUUGA	7885
2179	ACGUUAAU A UGGGCCUA	486	UAGGCCCA CUGAUGAG GCCGUUAGGC CGAA AUUAACGU	7886
2187	AUGGGCCU A AAAAUCAG	487	CUGAUUUU CUGAUGAG GCCGUUAGGC CGAA AGGCCCAU	7887

2193	CUAAAAAU C AGACAACU	488	AGUUGUCU CUGAUGAG GCCGUUAGGC CGAA AUUUUUAG	7888
2202	AGACAACU A UUGUGGUU	489	AACCACAA CUGAUGAG GCCGUUAGGC CGAA AGUUGUCU	7889
2204	ACAACUAU U GUGGUUUC	490	GAAACCAC CUGAUGAG GCCGUUAGGC CGAA AUAGUUGU	7890
2210	AUUGUGGU U UCACAUUU	491	AAAUGUGA CUGAUGAG GCCGUUAGGC CGAA ACCACAAU	7891
2211	UUGUGGUU U CACAUUUC	492	GAAAUGUG CUGAUGAG GCCGUUAGGC CGAA AACCACAA	7892
2212	UGUGGUUU C ACAUUUCC	493	GGAAAUGU CUGAUGAG GCCGUUAGGC CGAA AAACCACA	7893
2217	UUUCACAU U UCCUGUCU	494	AGACAGGA CUGAUGAG GCCGUUAGGC CGAA AUGUGAAA	7894
2218	UUCACAUU U CCUGUCUU	495	AAGACAGG CUGAUGAG GCCGUUAGGC CGAA AAUGUGAA	7895
2219	UCACAUUU C CUGUCUUA	496	UAAGACAG CUGAUGAG GCCGUUAGGC CGAA AAAUGUGA	7896
2224	UUUCCUGU C UUACUUUU	497	AAAAGUAA CUGAUGAG GCCGUUAGGC CGAA ACAGGAAA	7897
2226	UCCUGUCU U ACUUUUGG	498	CCAAAAGU CUGAUGAG GCCGUUAGGC CGAA AGACAGGA	7898
2227	CCUGUCUU A CUUUUGGG	499	CCCAAAAG CUGAUGAG GCCGUUAGGC CGAA AAGACAGG	7899
2230	GUCUUACU U UUGGGCGA	500	UCGCCCAA CUGAUGAG GCCGUUAGGC CGAA AGUAAGAC	7900
2231	UCUUACUU U UGGGCGAG	501	CUCGCCCA CUGAUGAG GCCGUUAGGC CGAA AAGUAAGA	7901
2232	CUUACUUU U GGGCGAGA	502	UCUCGCCC CUGAUGAG GCCGUUAGGC CGAA AAAGUAAG	7902
2247	GAAACUGU U CUUGAAUA	503	UAUUCAAG CUGAUGAG GCCGUUAGGC CGAA ACAGUUUC	7903
2248	AAACUGUU C UUGAAUAU	504	AUAUUCAA CUGAUGAG GCCGUUAGGC CGAA AACAGUUU	7904
2250	ACUGUUCU U GAAUAUUU	505	AAAUAUUC CUGAUGAG GCCGUUAGGC CGAA AGAACAGU	7905
2255	UCUUGAAU A UUUGGUGU	506	ACACCAAA CUGAUGAG GCCGUUAGGC CGAA AUUCAAGA	7906
2257	UUGAAUAU U UGGUGUCU	507	AGACACCA CUGAUGAG GCCGUUAGGC CGAA AUAUUCAA	7907
2258	UGAAUAUU U GGUGUCUU	508	AAGACACC CUGAUGAG GCCGUUAGGC CGAA AAUAUUCA	7908
2264	UUUGGUGU C UUUUGGAG	509	CUCCAAAA CUGAUGAG GCCGUUAGGC CGAA ACACCAAA	7909
2266	UGGUGUCU U UUGGAGUG	510	CACUCCAA CUGAUGAG GCCGUUAGGC CGAA AGACACCA	7910
2267	GGUGUCUU U UGGAGUGU	511	ACACUCCA CUGAUGAG GCCGUUAGGC CGAA AAGACACC	7911
2268	GUGUCUUU U GGAGUGUG	512	CACACUCC CUGAUGAG GCCGUUAGGC CGAA AAAGACAC	7912
2280	GUGUGGAU U CGCACUCC	513	GGAGUGCG CUGAUGAG GCCGUUAGGC CGAA AUCCACAC	7913
2281	UGUGGAUU C GCACUCCU	514	AGGAGUGC CUGAUGAG GCCGUUAGGC CGAA AAUCCACA	7914
2287	UUCGCACU C CUCCUGCA	515	UGCAGGAG CUGAUGAG GCCGUUAGGC CGAA AGUGCGAA	7915
2290	GCACUCCU C CUGCAUAU	516	AUAUGCAG CUGAUGAG GCCGUUAGGC CGAA AGGAGUGC	7916
2297	UCCUGCAU A UAGACCAC	517	GUGGUCUA CUGAUGAG GCCGUUAGGC CGAA AUGCAGGA	7917
2299	CUGCAUAU A GACCACCA	518	UGGUGGUC CUGAUGAG GCCGUUAGGC CGAA AUAUGCAG	7918
2317	AUGCCCCU A UCUUAUCA	519	UGAUAAGA CUGAUGAG GCCGUUAGGC CGAA AGGGGCAU	7919
2319	GCCCCUAU C UUAUCAAC	520	GUUGAUAA CUGAUGAG GCCGUUAGGC CGAA AUAGGGGC	7920
2321	CCCUAUCU U AUCAACAC	521	GUGUUGAU CUGAUGAG GCCGUUAGGC CGAA AGAUAGGG	7921
2322	CCUAUCUU A UCAACACU	522	AGUGUUGA CUGAUGAG GCCGUUAGGC CGAA AAGAUAGG	7922
2324	UAUCUUAU C AACACUUC	523	GAAGUGUU CUGAUGAG GCCGUUAGGC CGAA AUAAGAUA	7923
2331	UCAACACU U CCGGAAAC	524	GUUUCCGG CUGAUGAG GCCGUUAGGC CGAA AGUGUUGA	7924
2332	CAACACUU C CGGAAACU	525	AGUUUCCG CUGAUGAG GCCGUUAGGC CGAA AAGUGUUG	7925
2341	CGGAAACU A CUGUUGUU	526	AACAACAG CUGAUGAG GCCGUUAGGC CGAA AGUUUCCG	7926
2346	ACUACUGU U GUUAGACG	527	CGUCUAAC CUGAUGAG GCCGUUAGGC CGAA ACAGUAGU	7927
2349	ACUGUUGU U AGACGAAG	528	CUUCGUCU CUGAUGAG GCCGUUAGGC CGAA ACAACAGU	7928
2350	CUGUUGUU A GACGAAGA	529	UCUUCGUC CUGAUGAG GCCGUUAGGC CGAA AACAACAG	7929
2366	AGGCAGGU C CCCUAGAA	530	UUCUAGGG CUGAUGAG GCCGUUAGGC CGAA ACCUGCCU	7930
2371	GGUCCCCU A GAAGAAGA	531	UCUUCUUC CUGAUGAG GCCGUUAGGC CGAA AGGGGACC	7931
2383	GAAGAACU C CCUCGCCU	532	AGGCGAGG CUGAUGAG GCCGUUAGGC CGAA AGUUCUUC	7932
2387	AACUCCCU C GCCUCGCA	533	UGCGAGGC CUGAUGAG GCCGUUAGGC CGAA AGGGAGUU	7933
2392	CCUCGCCU C GCAGACGA	534	UCGUCUGC CUGAUGAG GCCGUUAGGC CGAA AGGCGAGG	7934
2405	ACGAAGGU C UCAAUCGC	535	GCGAUUGA CUGAUGAG GCCGUUAGGC CGAA ACCUUCGU	7935
2407	GAAGGUCU C AAUCGCCG	536	CGGCGAUU CUGAUGAG GCCGUUAGGC CGAA AGACCUUC	7936
2411	GUCUCAAU C GCCGCGUC	537	GACGCGGC CUGAUGAG GCCGUUAGGC CGAA AUUGAGAC	7937
2419	CGCCGCGU C GCAGAAGA	538	UCUUCUGC CUGAUGAG GCCGUUAGGC CGAA ACGCGGCG	7938

2429	CAGAAGAU C UCAAUCUC	539	GAGAUUGA CUGAUGAG GCCGUUAGGC CGAA AUCUUCUG	7939
2431	GAAGAUCU C AAUCUCGG	540	CCGAGAUU CUGAUGAG GCCGUUAGGC CGAA AGAUCUUC	7940
2435	AUCUCAAU C UCGGGAAU	541	AUUCCCGA CUGAUGAG GCCGUUAGGC CGAA AUUGAGAU	7941
2437	CUCAAUCU C GGGAAUCU	542	AGAUUCCC CUGAUGAG GCCGUUAGGC CGAA AGAUUGAG	7942
2444	UCGGGAAU C UCAAUGUU	543	AACAUUGA CUGAUGAG GCCGUUAGGC CGAA AUUCCCGA	7943
2446	GGGAAUCU C AAUGUUAG	544	CUAACAUU CUGAUGAG GCCGUUAGGC CGAA AGAUUCCC	7944
2452	CUCAAUGU U AGUAUUCC	545	GGAAUACU CUGAUGAG GCCGUUAGGC CGAA ACAUUGAG	7945
2453	UCAAUGUU A GUAUUCCU	546	AGGAAUAC CUGAUGAG GCCGUUAGGC CGAA AACAUUGA	7946
2456	AUGUUAGU A UUCCUUGG	547	CCAAGGAA CUGAUGAG GCCGUUAGGC CGAA ACUAACAU	7947
2458	GUUAGUAU U CCUUGGAC	548	GUCCAAGG CUGAUGAG GCCGUUAGGC CGAA AUACUAAC	7948
2459	UUAGUAUU C CUUGGACA	549	UGUCCAAG CUGAUGAG GCCGUUAGGC CGAA AAUACUAA	7949
2462	GUAUUCCU U GGACACAU	550	AUGUGUCC CUGAUGAG GCCGUUAGGC CGAA AGGAAUAC	7950
2471	GGACACAU A AGGUGGGA	551	UCCCACCU CUGAUGAG GCCGUUAGGC CGAA AUGUGUCC	7951
2484	GGGAAACU U UACGGGGC	552	GCCCCGUA CUGAUGAG GCCGUUAGGC CGAA AGUUUCCC	7952
2485	GGAAACUU U ACGGGGCU	553	AGCCCCGU CUGAUGAG GCCGUUAGGC CGAA AAGUUUCC	7953
2486	GAAACUUU A CGGGGCUU	554	AAGCCCCG CUGAUGAG GCCGUUAGGC CGAA AAAGUUUC	7954
2494	ACGGGGCU U UAUUCUUC	555	GAAGAAUA CUGAUGAG GCCGUUAGGC CGAA AGCCCCGU	7955
2495	CGGGGCUU U AUUCUUCU	556	AGAAGAAU CUGAUGAG GCCGUUAGGC CGAA AAGCCCCG	7956
2496	GGGGCUUU A UUCUUCUA	557	UAGAAGAA CUGAUGAG GCCGUUAGGC CGAA AAAGCCCC	7957
2498	GGCUUUAU U CUUCUACG	558	CGUAGAAG CUGAUGAG GCCGUUAGGC CGAA AUAAAGCC	7958
2499	GCUUUAUU C UUCUACGG	559	CCGUAGAA CUGAUGAG GCCGUUAGGC CGAA AAUAAAGC	7959
2501	UUUAUUCU U CUACGGUA	560	UACCGUAG CUGAUGAG GCCGUUAGGC CGAA AGAAUAAA	7960
2502	UUAUUCUU C UACGGUAC	561	GUACCGUA CUGAUGAG GCCGUUAGGC CGAA AAGAAUAA	7961
2504	AUUCUUCU A CGGUACCU	562	AGGUACCG CUGAUGAG GCCGUUAGGC CGAA AGAAGAAU	7962
2509	UCUACGGU A CCUUGCUU	563	AAGCAAGG CUGAUGAG GCCGUUAGGC CGAA ACCGUAGA	7963
2513	CGGUACCU U GCUUUAAU	564	AUUAAAGC CUGAUGAG GCCGUUAGGC CGAA AGGUACCG	7964
2517	ACCUUGCU U UAAUCCUA	565	UAGGAUUA CUGAUGAG GCCGUUAGGC CGAA AGCAAGGU	7965
2518	CCUUGCUU U AAUCCUAA	566	UUAGGAUU CUGAUGAG GCCGUUAGGC CGAA AAGCAAGG	7966
2519	CUUGCUUU A AUCCUAAA	567	UUUAGGAU CUGAUGAG GCCGUUAGGC CGAA AAAGCAAG	7967
2522	GCUUUAAU C CUAAAUGG	568	CCAUUUAG CUGAUGAG GCCGUUAGGC CGAA AUUAAAGC	7968
2525	UUAAUCCU A AAUGGCAA	569	UUGCCAUU CUGAUGAG GCCGUUAGGC CGAA AGGAUUAA	7969
2537	GGCAAACU C CUUCUUUU	570	AAAAGAAG CUGAUGAG GCCGUUAGGC CGAA AGUUUGCC	7970
2540	AAACUCCU U CUUUUCCU	571	AGGAAAAG CUGAUGAG GCCGUUAGGC CGAA AGGAGUUU	7971
2541	AACUCCUU C UUUUCCUG	572	CAGGAAAA CUGAUGAG GCCGUUAGGC CGAA AAGGAGUU	7972
2543	CUCCUUCU U UUCCUGAC	573	GUCAGGAA CUGAUGAG GCCGUUAGGC CGAA AGAAGGAG	7973
2544	UCCUUCUU U UCCUGACA	574	UGUCAGGA CUGAUGAG GCCGUUAGGC CGAA AAGAAGGA	7974
2545	CCUUCUUU U CCUGACAU	575	AUGUCAGG CUGAUGAG GCCGUUAGGC CGAA AAAGAAGG	7975
2546	CUUCUUUU C CUGACAUU	576	AAUGUCAG CUGAUGAG GCCGUUAGGC CGAA AAAAGAAG	7976
2554	CCUGACAU U CAUUUGCA	577	UGCAAAUG CUGAUGAG GCCGUUAGGC CGAA AUGUCAGG	7977
2555	CUGACAUU C AUUUGCAG	578	CUGCAAAU CUGAUGAG GCCGUUAGGC CGAA AAUGUCAG	7978
2558	ACAUUCAU U UGCAGGAG	579	CUCCUGCA CUGAUGAG GCCGUUAGGC CGAA AUGAAUGU	7979
2559	CAUUCAUU U GCAGGAGG	580	CCUCCUGC CUGAUGAG GCCGUUAGGC CGAA AAUGAAUG	7980
2572	GAGGACAU U GUUGAUAG	581	CUAUCAAC CUGAUGAG GCCGUUAGGC CGAA AUGUCCUC	7981
2575	GACAUUGU U GAUAGAUG	582	CAUCUAUC CUGAUGAG GCCGUUAGGC CGAA ACAAUGUC	7982
2579	UUGUUGAU A GAUGUAAG	583	CUUACAUC CUGAUGAG GCCGUUAGGC CGAA AUCAACAA	7983
2585	AUAGAUGU A AGCAAUUU	584	AAAUUGCU CUGAUGAG GCCGUUAGGC CGAA ACAUCUAU	7984
2592	UAAGCAAU U UGUGGGGC	585	GCCCCACA CUGAUGAG GCCGUUAGGC CGAA AUUGCUUA	7985
2593	AAGCAAUU U GUGGGGCC	586	GGCCCCAC CUGAUGAG GCCGUUAGGC CGAA AAUUGCUU	7986
2605	GGGCCCCU U ACAGUAAA	587	UUUACUGU CUGAUGAG GCCGUUAGGC CGAA AGGGGCCC	7987
2606	GGCCCCUU A CAGUAAAU	588	AUUUACUG CUGAUGAG GCCGUUAGGC CGAA AAGGGGCC	7988
2611	CUUACAGU A AAUGAAAA	589	UUUUCAUU CUGAUGAG GCCGUUAGGC CGAA ACUGUAAG	7989

2629	AGGAGACU U AAAUUAAC	590	GUUAAUUU CUGAUGAG GCCGUUAGGC CGAA AGUCUCCU	7990
2630	GGAGACUU A AAUUAACU	591	AGUUAAUU CUGAUGAG GCCGUUAGGC CGAA AAGUCUCC	7991
2634	ACUUAAAU U AACUAUGC	592	GCAUAGUU CUGAUGAG GCCGUUAGGC CGAA AUUUAAGU	7992
2635	CUUAAAUU A ACUAUGCC	593	GGCAUAGU CUGAUGAG GCCGUUAGGC CGAA AAUUUAAG	7993
2639	AAUUAACU A UGCCUGCU	594	AGCAGGCA CUGAUGAG GCCGUUAGGC CGAA AGUUAAUU	7994
2648	UGCCUGCU A GGUUUUAU	595	AUAAAACC CUGAUGAG GCCGUUAGGC CGAA AGCAGGCA	7995
2652	UGCUAGGU U UUAUCCCA	596	UGGGAUAA CUGAUGAG GCCGUUAGGC CGAA ACCUAGCA	7996
2653	GCUAGGUU U UAUCCCAA	597	UUGGGAUA CUGAUGAG GCCGUUAGGC CGAA AACCUAGC	7997
2654	CUAGGUUU U AUCCCAAU	598	AUUGGGAU CUGAUGAG GCCGUUAGGC CGAA AAACCUAG	7998
2655	UAGGUUUU A UCCCAAUG	599	CAUUGGGA CUGAUGAG GCCGUUAGGC CGAA AAAACCUA	7999
2657	GGUUUUAU C CCAAUGUU	600	AACAUUGG CUGAUGAG GCCGUUAGGC CGAA AUAAAACC	8000
2665	CCCAAUGU U ACUAAAUA	601	UAUUUAGU CUGAUGAG GCCGUUAGGC CGAA ACAUUGGG	8001
2666	CCAAUGUU A CUAAAUAU	602	AUAUUUAG CUGAUGAG GCCGUUAGGC CGAA AACAUUGG	8002
2669	AUGUUACU A AAUAUUUG	603	CAAAUAUU CUGAUGAG GCCGUUAGGC CGAA AGUAACAU	8003
2673	UACUAAAU A UUUGCCCU	604	AGGGCAAA CUGAUGAG GCCGUUAGGC CGAA AUUUAGUA	8004
2675	CUAAAUAU U UGCCCUUA	605	UAAGGGCA CUGAUGAG GCCGUUAGGC CGAA AUAUUUAG	8005
2676	UAAAUAUU U GCCCUUAG	606	CUAAGGGC CUGAUGAG GCCGUUAGGC CGAA AAUAUUUA	8006
2682	UUUGCCCU U AGAUAAAG	607	CUUUAUCU CUGAUGAG GCCGUUAGGC CGAA AGGGCAAA	8007
2683	UUGCCCUU A GAUAAAGG	608	CCUUUAUC CUGAUGAG GCCGUUAGGC CGAA AAGGGCAA	8008
2687	CCUUAGAU A AAGGGAUC	609	GAUCCCUU CUGAUGAG GCCGUUAGGC CGAA AUCUAAGG	8009
2695	AAAGGGAU C AAACCGUA	610	UACGGUUU CUGAUGAG GCCGUUAGGC CGAA AUCCCUUU	8010
2703	CAAACCGU A UUAUCCAG	611	CUGGAUAA CUGAUGAG GCCGUUAGGC CGAA ACGGUUUG	8011
2705	AACCGUAU U AUCCAGAG	612	CUCUGGAU CUGAUGAG GCCGUUAGGC CGAA AUACGGUU	8012
2706	ACCGUAUU A UCCAGAGU	613	ACUCUGGA CUGAUGAG GCCGUUAGGC CGAA AAUACGGU	8013
2708	CGUAUUAU C CAGAGUAU	614	AUACUCUG CUGAUGAG GCCGUUAGGC CGAA AUAAUACG	8014
2715	UCCAGAGU A UGUAGUUA	615	UAACUACA CUGAUGAG GCCGUUAGGC CGAA ACUCUGGA	8015
2719	GAGUAUGU A GUUAAUCA	616	UGAUUAAC CUGAUGAG GCCGUUAGGC CGAA ACAUACUC	8016
2722	UAUGUAGU U AAUCAUUA	617	UAAUGAUU CUGAUGAG GCCGUUAGGC CGAA ACUACAUA	8017
2723	AUGUAGUU A AUCAUUAC	618	GUAAUGAU CUGAUGAG GCCGUUAGGC CGAA AACUACAU	8018
2726	UAGUUAAU C AUUACUUC	619	GAAGUAAU CUGAUGAG GCCGUUAGGC CGAA AUUAACUA	8019
2729	UUAAUCAU U ACUUCCAG	620	CUGGAAGU CUGAUGAG GCCGUUAGGC CGAA AUGAUUAA	8020
2730	UAAUCAUU A CUUCCAGA	621	UCUGGAAG CUGAUGAG GCCGUUAGGC CGAA AAUGAUUA	8021
2733	UCAUUACU U CCAGACGC	622	GCGUCUGG CUGAUGAG GCCGUUAGGC CGAA AGUAAUGA	8022
2734	CAUUACUU C CAGACGCG	623	CGCGUCUG CUGAUGAG GCCGUUAGGC CGAA AAGUAAUG	8023
2747	CGCGACAU U AUUUACAC	624	GUGUAAAU CUGAUGAG GCCGUUAGGC CGAA AUGUCGCG	8024
2748	GCGACAUU A UUUACACA	625	UGUGUAAA CUGAUGAG GCCGUUAGGC CGAA AAUGUCGC	8025
2750	GACAUUAU U UACACACU	626	AGUGUGUA CUGAUGAG GCCGUUAGGC CGAA AUAAUGUC	8026
2751	ACAUUAUU U ACACACUC	627	GAGUGUGU CUGAUGAG GCCGUUAGGC CGAA AAUAAUGU	8027
2752	CAUUAUUU A CACACUCU	628	AGAGUGUG CUGAUGAG GCCGUUAGGC CGAA AAAUAAUG	8028
2759	UACACACU C UUUGGAAG	629	CUUCCAAA CUGAUGAG GCCGUUAGGC CGAA AGUGUGUA	8029
2761	CACACUCU U UGGAAGGC	630	GCCUUCCA CUGAUGAG GCCGUUAGGC CGAA AGAGUGUG	8030
2762	ACACUCUU U GGAAGGCG	631	CGCCUUCC CUGAUGAG GCCGUUAGGC CGAA AAGAGUGU	8031
2776	GCGGGGAU C UUAUAUAA	632	UUAUAUAA CUGAUGAG GCCGUUAGGC CGAA AUCCCCGC	8032
2778	GGGGAUCU U AUAUAAAA	633	UUUUAUAU CUGAUGAG GCCGUUAGGC CGAA AGAUCCCC	8033
2779	GGGAUCUU A UAUAAAAG	634	CUUUUAUA CUGAUGAG GCCGUUAGGC CGAA AAGAUCCC	8034
2781	GAUCUUAU A UAAAAGAG	635	CUCUUUUA CUGAUGAG GCCGUUAGGC CGAA AUAAGAUC	8035
2783	UCUUAUAU A AAAGAGAG	636	CUCUCUUU CUGAUGAG GCCGUUAGGC CGAA AUAUAAGA	8036
2793	AAGAGAGU C CACACGUA	637	UACGUGUG CUGAUGAG GCCGUUAGGC CGAA ACUCUCUU	8037
2801	CCACACGU A GCGCCUCA	638	UGAGGCGC CUGAUGAG GCCGUUAGGC CGAA ACGUGUGG	8038
2808	UAGCGCCU C AUUUUGCG	639	CGCAAAAU CUGAUGAG GCCGUUAGGC CGAA AGGCGCUA	8039
2811	CGCCUCAU U UUGCGGGU	640	ACCCGCAA CUGAUGAG GCCGUUAGGC CGAA AUGAGGCG	8040

2812	GCCUCAUU U UGCGGGUC	641	GACCCGCA CUGAUGAG GCCGUUAGGC CGAA AAUGAGGC	8041
2813	CCUCAUUU U GCGGGUCA	642	UGACCCGC CUGAUGAG GCCGUUAGGC CGAA AAAUGAGG	8042
2820	UUGCGGGU C ACCAUAUU	643	AAUAUGGU CUGAUGAG GCCGUUAGGC CGAA ACCCGCAA	8043
2826	GUCACCAU A UUCUUGGG	644	CCCAAGAA CUGAUGAG GCCGUUAGGC CGAA AUGGUGAC	8044
2828	CACCAUAU U CUUGGGAA	645	UUCCCAAG CUGAUGAG GCCGUUAGGC CGAA AUAUGGUG	8045
2829	ACCAUAUU C UUGGGAAC	646	GUUCCCAA CUGAUGAG GCCGUUAGGC CGAA AAUAUGGU	8046
2831	CAUAUUCU U GGGAACAA	647	UUGUUCCC CUGAUGAG GCCGUUAGGC CGAA AGAAUAUG	8047
2843	AACAAGAU C UACAGCAU	648	AUGCUGUA CUGAUGAG GCCGUUAGGC CGAA AUCUUGUU	8048
2845	CAAGAUCU A CAGCAUGG	649	CCAUGCUG CUGAUGAG GCCGUUAGGC CGAA AGAUCUUG	8049
2859	UGGGAGGU U GGUCUUCC	650	GGAAGACC CUGAUGAG GCCGUUAGGC CGAA ACCUCCCA	8050
2863	AGGUUGGU C UUCCAAAC	651	GUUUGGAA CUGAUGAG GCCGUUAGGC CGAA ACCAACCU	8051
2865	GUUGGUCU U CCAAACCU	652	AGGUUUGG CUGAUGAG GCCGUUAGGC CGAA AGACCAAC	8052
2866	UUGGUCUU C CAAACCUC	653	GAGGUUUG CUGAUGAG GCCGUUAGGC CGAA AAGACCAA	8053
2874	CCAAACCU C GAAAAGGC	654	GCCUUUUC CUGAUGAG GCCGUUAGGC CGAA AGGUUUGG	8054
2895	GGACAAAU C UUUCUGUC	655	GACAGAAA CUGAUGAG GCCGUUAGGC CGAA AUUUGUCC	8055
2897	ACAAAUCU U UCUGUCCC	656	GGGACAGA CUGAUGAG GCCGUUAGGC CGAA AGAUUUGU	8056
2898	CAAAUCUU U CUGUCCCC	657	GGGGACAG CUGAUGAG GCCGUUAGGC CGAA AAGAUUUG	8057
2899	AAAUCUUU C UGUCCCCA	658	UGGGGACA CUGAUGAG GCCGUUAGGC CGAA AAAGAUUU	8058
2903	CUUUCUGU C CCCAAUCC	659	GGAUUGGG CUGAUGAG GCCGUUAGGC CGAA ACAGAAAG	8059
2910	UCCCCAAU C CCCUGGGA	660	UCCCAGGG CUGAUGAG GCCGUUAGGC CGAA AUUGGGGA	8060
2920	CCUGGGAU U CUUCCCCG	661	CGGGGAAG CUGAUGAG GCCGUUAGGC CGAA AUCCCAGG	8061
2921	CUGGGAUU C UUCCCCGA	662	UCGGGGAA CUGAUGAG GCCGUUAGGC CGAA AAUCCCAG	8062
2923	GGGAUUCU U CCCCGAUC	663	GAUCGGGG CUGAUGAG GCCGUUAGGC CGAA AGAAUCCC	8063
2924	GGAUUCUU C CCCGAUCA	664	UGAUCGGG CUGAUGAG GCCGUUAGGC CGAA AAGAAUCC	8064
2931	UCCCCGAU C AUCAGUUG	665	CAACUGAU CUGAUGAG GCCGUUAGGC CGAA AUCGGGGA	8065
2934	CCGAUCAU C AGUUGGAC	666	GUCCAACU CUGAUGAG GCCGUUAGGC CGAA AUGAUCGG	8066
2938	UCAUCAGU U GGACCCUG	667	CAGGGUCC CUGAUGAG GCCGUUAGGC CGAA ACUGAUGA	8067
2950	CCCUGCAU U CAAAGCCA	668	UGGCUUUG CUGAUGAG GCCGUUAGGC CGAA AUGCAGGG	8068
2951	CCUGCAUU C AAAGCCAA	669	UUGGCUUU CUGAUGAG GCCGUUAGGC CGAA AAUGCAGG	8069
2962	AGCCAACU C AGUAAAUC	670	GAUUUACU CUGAUGAG GCCGUUAGGC CGAA AGUUGGCU	8070
2966	AACUCAGU A AAUCCAGA	671	UCUGGAUU CUGAUGAG GCCGUUAGGC CGAA ACUGAGUU	8071
2970	CAGUAAAU C CAGAUUGG	672	CCAAUCUG CUGAUGAG GCCGUUAGGC CGAA AUUUACUG	8072
2976	AUCCAGAU U GGGACCUC	673	GAGGUCCC CUGAUGAG GCCGUUAGGC CGAA AUCUGGAU	8073
2984	UGGGACCU C AACCCGCA	674	UGCGGGUU CUGAUGAG GCCGUUAGGC CGAA AGGUCCCA	8074
3037	GGGAGCAU U CGGGCCAG	675	CUGGCCCG CUGAUGAG GCCGUUAGGC CGAA AUGCUCCC	8075
3038	GGAGCAUU C GGGCCAGG	676	CCUGGCCC CUGAUGAG GCCGUUAGGC CGAA AAUGCUCC	8076
3049	GCCAGGGU U CACCCCUC	677	GAGGGGUG CUGAUGAG GCCGUUAGGC CGAA ACCCUGGC	8077
3050	CCAGGGUU C ACCCCUCC	678	GGAGGGGU CUGAUGAG GCCGUUAGGC CGAA AACCCUGG	8078
3057	UCACCCCU C CCCAUGGG	679	CCCAUGGG CUGAUGAG GCCGUUAGGC CGAA AGGGGUGA	8079
3073	GGGACUGU U GGGGUGGA	680	UCCACCCC CUGAUGAG GCCGUUAGGC CGAA ACAGUCCC	8080
3087	GGAGCCCU C ACGCUCAG	681	CUGAGCGU CUGAUGAG GCCGUUAGGC CGAA AGGGCUCC	8081
3093	CUCACGCU C AGGGCCUA	682	UAGGCCCU CUGAUGAG GCCGUUAGGC CGAA AGCGUGAG	8082
3101	CAGGGCCU A CUCACAAC	683	GUUGUGAG CUGAUGAG GCCGUUAGGC CGAA AGGCCCUG	8083
3104	GGCCUACU C ACAACUGU	684	ACAGUUGU CUGAUGAG GCCGUUAGGC CGAA AGUAGGCC	8084
3123	CAGCAGCU C CUCCUCCU	685	AGGAGGAG CUGAUGAG GCCGUUAGGC CGAA AGCUGCUG	8085
3126	CAGCUCCU C CUCCUGCC	686	GGCAGGAG CUGAUGAG GCCGUUAGGC CGAA AGGAGCUG	8086
3129	CUCCUCCU C CUGCCUCC	687	GGAGGCAG CUGAUGAG GCCGUUAGGC CGAA AGGAGGAG	8087
3136	UCCUGCCU C CACCAAUC	688	GAUUGGUG CUGAUGAG GCCGUUAGGC CGAA AGGCAGGA	8088
3144	CCACCAAU C GGCAGUCA	689	UGACUGCC CUGAUGAG GCCGUUAGGC CGAA AUUGGUGG	8089
3151	UCGGCAGU C AGGAAGGC	690	GCCUUCCU CUGAUGAG GCCGUUAGGC CGAA ACUGCCGA	8090
3165	GGCAGCCU A CUCCCUUA	691	UAAGGGAG CUGAUGAG GCCGUUAGGC CGAA AGGCUGCC	8091

3168	AGCCUACU C CCUUAUCU	692	AGAUAAGG CUGAUGAG GCCGUUAGGC CGAA AGUAGGCU	8092
3172	UACUCCCU U AUCUCCAC	693	GUGGAGAU CUGAUGAG GCCGUUAGGC CGAA AGGGAGUA	8093
3173	ACUCCCUU A UCUCCACC	694	GGUGGAGA CUGAUGAG GCCGUUAGGC CGAA AAGGGAGU	8094
3175	UCCCUUAU C UCCACCUC	695	GAGGUGGA CUGAUGAG GCCGUUAGGC CGAA AUAAGGGA	8095
3177	CCUUAUCU C CACCUCUA	696	UAGAGGUG CUGAUGAG GCCGUUAGGC CGAA AGAUAAGG	8096
3183	CUCCACCU C UAAGGGAC	697	GUCCCUUA CUGAUGAG GCCGUUAGGC CGAA AGGUGGAG	8097
3185	CCACCUCU A AGGGACAC	698	GUGUCCCU CUGAUGAG GCCGUUAGGC CGAA AGAGGUGG	8098
3195	GGGACACU C AUCCUCAG	699	CUGAGGAU CUGAUGAG GCCGUUAGGC CGAA AGUGUCCC	8099
3198	ACACUCAU C CUCAGGCC	700	GGCCUGAG CUGAUGAG GCCGUUAGGC CGAA AUGAGUGU	8100
3201	CUCAUCCU C AGGCCAUG	701	CAUGGCCU CUGAUGAG GCCGUUAGGC CGAA AGGAUGAG	8101

Input Sequence = AF100308. Cut Site = UH/.
Stem Length = 8 . Core Sequence = CUGAUGAG GCCGUUAGGC CGAA
AF100308 (Hepatitis B virus strain 2-18, 3215 bp)

Underlined region can be any X sequence or linker, as described herein.

TABLE VI: HUMAN HBV INOZYME AND SUBSTRATE SEQUENCE

Pos	Substrate	Seq ID	Inozyme	Seq ID
9	AACUCCAC C ACUUUCCA	702	UGGAAAGU CUGAUGAG GCCGUUAGGC CGAA IUGGAGUU	8102
10	ACUCCACC A CUUUCCAC	703	GUGGAAAG CUGAUGAG GCCGUUAGGC CGAA IGUGGAGU	8103
12	UCCACCAC U UUCCACCA	704	UGGUGGAA CUGAUGAG GCCGUUAGGC CGAA IUGGUGGA	8104
16	CCACUUUC C ACCAAACU	705	AGUUUGGU CUGAUGAG GCCGUUAGGC CGAA IAAAGUGG	8105
17	CACUUUCC A CCAAACUC	706	GAGUUUGG CUGAUGAG GCCGUUAGGC CGAA IGAAAGUG	8106
19	CUUUCCAC C AAACUCUU	707	AAGAGUUU CUGAUGAG GCCGUUAGGC CGAA IUGGAAAG	8107
20	UUUCCACC A AACUCUUC	708	GAAGAGUU CUGAUGAG GCCGUUAGGC CGAA IGUGGAAA	8108
24	CACCAAAC U CUUCAAGA	709	UCUUGAAG CUGAUGAG GCCGUUAGGC CGAA IUUUGGUG	8109
26	CCAAACUC U UCAAGAUC	710	GAUCUUGA CUGAUGAG GCCGUUAGGC CGAA IAGUUUGG	8110
29	AACUCUUC A AGAUCCCA	711	UGGGAUCU CUGAUGAG GCCGUUAGGC CGAA IAAGAGUU	8111
35	UCAAGAUC C CAGAGUCA	712	UGACUCUG CUGAUGAG GCCGUUAGGC CGAA IAUCUUGA	8112
36	CAAGAUCC C AGAGUCAG	713	CUGACUCU CUGAUGAG GCCGUUAGGC CGAA IGAUCUUG	8113
37	AAGAUCCC A GAGUCAGG	714	CCUGACUC CUGAUGAG GCCGUUAGGC CGAA IGGAUCUU	8114
43	CCAGAGUC A GGGCCCUG	715	CAGGGCCC CUGAUGAG GCCGUUAGGC CGAA IACUCUGG	8115
48	GUCAGGGC C CUGUACUU	716	AAGUACAG CUGAUGAG GCCGUUAGGC CGAA ICCCUGAC	8116
49	UCAGGGCC C UGUACUUU	717	AAAGUACA CUGAUGAG GCCGUUAGGC CGAA IGCCCUGA	8117
50	CAGGGCCC U GUACUUUC	718	GAAAGUAC CUGAUGAG GCCGUUAGGC CGAA IGGCCCUG	8118
55	CCCUGUAC U UUCCUGCU	719	AGCAGGAA CUGAUGAG GCCGUUAGGC CGAA IUACAGGG	8119
59	GUACUUUC C UGCUGGUG	720	CACCAGCA CUGAUGAG GCCGUUAGGC CGAA IAAAGUAC	8120
60	UACUUUCC U GCUGGUGG	721	CCACCAGC CUGAUGAG GCCGUUAGGC CGAA IGAAAGUA	8121
63	UUUCCUGC U GGUGGCUC	722	GAGCCACC CUGAUGAG GCCGUUAGGC CGAA ICAGGAAA	8122
70	CUGGUGGC U CCAGUUCA	723	UGAACUGG CUGAUGAG GCCGUUAGGC CGAA ICCACCAG	8123
72	GGUGGCUC C AGUUCAGG	724	CCUGAACU CUGAUGAG GCCGUUAGGC CGAA IAGCCACC	8124
73	GUGGCUCC A GUUCAGGA	725	UCCUGAAC CUGAUGAG GCCGUUAGGC CGAA IGAGCCAC	8125
78	UCCAGUUC A GGAACAGU	726	ACUGUUCC CUGAUGAG GCCGUUAGGC CGAA IAACUGGA	8126
84	UCAGGAAC A GUGAGCCC	727	GGGCUCAC CUGAUGAG GCCGUUAGGC CGAA IUUCCUGA	8127
91	CAGUGAGC C CUGCUCAG	728	CUGAGCAG CUGAUGAG GCCGUUAGGC CGAA ICUCACUG	8128
92	AGUGAGCC C UGCUCAGA	729	UCUGAGCA CUGAUGAG GCCGUUAGGC CGAA IGCUCACU	8129
93	GUGAGCCC U GCUCAGAA	730	UUCUGAGC CUGAUGAG GCCGUUAGGC CGAA IGGCUCAC	8130
96	AGCCCUGC U CAGAAUAC	731	GUAUUCUG CUGAUGAG GCCGUUAGGC CGAA ICAGGGCU	8131
98	CCCUGCUC A GAAUACUG	732	CAGUAUUC CUGAUGAG GCCGUUAGGC CGAA IAGCAGGG	8132
105	CAGAAUAC U GUCUCUGC	733	GCAGAGAC CUGAUGAG GCCGUUAGGC CGAA IUAUUCUG	8133
109	AUACUGUC U CUGCCAUA	734	UAUGGCAG CUGAUGAG GCCGUUAGGC CGAA IACAGUAU	8134
111	ACUGUCUC U GCCAUAUC	735	GAUAUGGC CUGAUGAG GCCGUUAGGC CGAA IAGACAGU	8135
114	GUCUCUGC C AUAUCGUC	736	GACGAUAU CUGAUGAG GCCGUUAGGC CGAA ICAGAGAC	8136
115	UCUCUGCC A UAUCGUCA	737	UGACGAUA CUGAUGAG GCCGUUAGGC CGAA IGCAGAGA	8137
123	AUAUCGUC A AUCUUAUC	738	GAUAAGAU CUGAUGAG GCCGUUAGGC CGAA IACGAUAU	8138
127	CGUCAAUC U UAUCGAAG	739	CUUCGAUA CUGAUGAG GCCGUUAGGC CGAA IAUUGACG	8139
138	UCGAAGAC U GGGGACCC	740	GGGUCCCC CUGAUGAG GCCGUUAGGC CGAA IUCUUCGA	8140
145	CUGGGGAC C CUGUACCG	741	CGGUACAG CUGAUGAG GCCGUUAGGC CGAA IUCCCCAG	8141
146	UGGGGACC C UGUACCGA	742	UCGGUACA CUGAUGAG GCCGUUAGGC CGAA IGUCCCCA	8142
147	GGGGACCC U GUACCGAA	743	UUCGGUAC CUGAUGAG GCCGUUAGGC CGAA IGGUCCCC	8143
152	CCCUGUAC C GAACAUGG	744	CCAUGUUC CUGAUGAG GCCGUUAGGC CGAA IUACAGGG	8144
157	UACCGAAC A UGGAGAAC	745	GUUCUCCA CUGAUGAG GCCGUUAGGC CGAA IUUCGGUA	8145
166	UGGAGAAC A UCGCAUCA	746	UGAUGCGA CUGAUGAG GCCGUUAGGC CGAA IUUCUCCA	8146
171	AACAUCGC A UCAGGACU	747	AGUCCUGA CUGAUGAG GCCGUUAGGC CGAA ICGAUGUU	8147

174	AUCGCAUC A GGACUCCU	748	AGGAGUCC CUGAUGAG GCCGUUAGGC CGAA IAUGCGAU	8148
179	AUCAGGAC U CCUAGGAC	749	GUCCUAGG CUGAUGAG GCCGUUAGGC CGAA IUCCUGAU	8149
181	CAGGACUC C UAGGACCC	750	GGGUCCUA CUGAUGAG GCCGUUAGGC CGAA IAGUCCUG	8150
182	AGGACUCC U AGGACCCC	751	GGGGUCCU CUGAUGAG GCCGUUAGGC CGAA IGAGUCCU	8151
188	CCUAGGAC C CCUGCUCG	752	CGAGCAGG CUGAUGAG GCCGUUAGGC CGAA IUCCUAGG	8152
189	CUAGGACC C CUGCUCGU	753	ACGAGCAG CUGAUGAG GCCGUUAGGC CGAA IGUCCUAG	8153
190	UAGGACCC C UGCUCGUG	754	CACGAGCA CUGAUGAG GCCGUUAGGC CGAA IGGUCCUA	8154
191	AGGACCCC U GCUCGUGU	755	ACACGAGC CUGAUGAG GCCGUUAGGC CGAA IGGGUCCU	8155
194	ACCCCUGC U CGUGUUAC	756	GUAACACG CUGAUGAG GCCGUUAGGC CGAA ICAGGGGU	8156
203	CGUGUUAC A GGCGGGGU	757	ACCCCGCC CUGAUGAG GCCGUUAGGC CGAA IUAACACG	8157
217	GGUUUUUC U UGUUGACA	758	UGUCAACA CUGAUGAG GCCGUUAGGC CGAA IAAAAACC	8158
225	UUGUUGAC A AAAAUCCU	759	AGGAUUUU CUGAUGAG GCCGUUAGGC CGAA IUCAACAA	8159
232	CAAAAAUC C UCACAAUA	760	UAUUGUGA CUGAUGAG GCCGUUAGGC CGAA IAUUUUUG	8160
233	AAAAAUCC U CACAAUAC	761	GUAUUGUG CUGAUGAG GCCGUUAGGC CGAA IGAUUUUU	8161
235	AAAUCCUC A CAAUACCA	762	UGGUAUUG CUGAUGAG GCCGUUAGGC CGAA IAGGAUUU	8162
237	AUCCUCAC A AUACCACA	763	UGUGGUAU CUGAUGAG GCCGUUAGGC CGAA IUGAGGAU	8163
242	CACAAUAC C ACAGAGUC	764	GACUCUGU CUGAUGAG GCCGUUAGGC CGAA IUAUUGUG	8164
243	ACAAUACC A CAGAGUCU	765	AGACUCUG CUGAUGAG GCCGUUAGGC CGAA IGUAUUGU	8165
245	AAUACCAC A GAGUCUAG	766	CUAGACUC CUGAUGAG GCCGUUAGGC CGAA IUGGUAUU	8166
251	ACAGAGUC U AGACUCGU	767	ACGAGUCU CUGAUGAG GCCGUUAGGC CGAA IACUCUGU	8167
256	GUCUAGAC U CGUGGUGG	768	CCACCACG CUGAUGAG GCCGUUAGGC CGAA IUCUAGAC	8168
267	UGGUGGAC U UCUCUCAA	769	UUGAGAGA CUGAUGAG GCCGUUAGGC CGAA IUCCACCA	8169
270	UGGACUUC U CUCAAUUU	770	AAAUUGAG CUGAUGAG GCCGUUAGGC CGAA IAAGUCCA	8170
272	GACUUCUC U CAAUUUUC	771	GAAAAUUG CUGAUGAG GCCGUUAGGC CGAA IAGAAGUC	8171
274	CUUCUCUC A AUUUUCUA	772	UAGAAAAU CUGAUGAG GCCGUUAGGC CGAA IAGAGAAG	8172
281	CAAUUUUC U AGGGGGAA	773	UUCCCCCU CUGAUGAG GCCGUUAGGC CGAA IAAAAUUG	8173
291	GGGGGAAC A CCCGUGUG	774	CACACGGG CUGAUGAG GCCGUUAGGC CGAA IUUCCCCC	8174
293	GGGAACAC C CGUGUGUC	775	GACACACG CUGAUGAG GCCGUUAGGC CGAA IUGUUCCC	8175
294	GGAACACC C GUGUGUCU	776	AGACACAC CUGAUGAG GCCGUUAGGC CGAA IGUGUUCC	8176
302	CGUGUGUC U UGGCCAAA	777	UUUGGCCA CUGAUGAG GCCGUUAGGC CGAA IACACACG	8177
307	GUCUUGGC C AAAAUUCG	778	CGAAUUUU CUGAUGAG GCCGUUAGGC CGAA ICCAAGAC	8178
308	UCUUGGCC A AAAUUCGC	779	GCGAAUUU CUGAUGAG GCCGUUAGGC CGAA IGCCAAGA	8179
317	AAAUUCGC A GUCCCAAA	780	UUUGGGAC CUGAUGAG GCCGUUAGGC CGAA ICGAAUUU	8180
321	UCGCAGUC C CAAAUCUC	781	GAGAUUUG CUGAUGAG GCCGUUAGGC CGAA IACUGCGA	8181
322	CGCAGUCC C AAAUCUCC	782	GGAGAUUU CUGAUGAG GCCGUUAGGC CGAA IGACUGCG	8182
323	GCAGUCCC A AAUCUCCA	783	UGGAGAUU CUGAUGAG GCCGUUAGGC CGAA IGGACUGC	8183
328	CCCAAAUC U CCAGUCAC	784	GUGACUGG CUGAUGAG GCCGUUAGGC CGAA IAUUUGGG	8184
330	CAAAUCUC C AGUCACUC	785	GAGUGACU CUGAUGAG GCCGUUAGGC CGAA IAGAUUUG	8185
331	AAAUCUCC A GUCACUCA	786	UGAGUGAC CUGAUGAG GCCGUUAGGC CGAA IGAGAUUU	8186
335	CUCCAGUC A CUCACCAA	787	UUGGUGAG CUGAUGAG GCCGUUAGGC CGAA IACUGGAG	8187
337	CCAGUCAC U CACCAACC	788	GGUUGGUG CUGAUGAG GCCGUUAGGC CGAA IUGACUGG	8188
339	AGUCACUC A CCAACCUG	789	CAGGUUGG CUGAUGAG GCCGUUAGGC CGAA IAGUGACU	8189
341	UCACUCAC C AACCUGUU	790	AACAGGUU CUGAUGAG GCCGUUAGGC CGAA IUGAGUGA	8190
342	CACUCACC A ACCUGUUG	791	CAACAGGU CUGAUGAG GCCGUUAGGC CGAA IGUGAGUG	8191
345	UCACCAAC C UGUUGUCC	792	GGACAACA CUGAUGAG GCCGUUAGGC CGAA IUUGGUGA	8192
346	CACCAACC U GUUGUCCU	793	AGGACAAC CUGAUGAG GCCGUUAGGC CGAA IGUUGGUG	8193
353	CUGUUGUC C UCCAAUUU	794	AAAUUGGA CUGAUGAG GCCGUUAGGC CGAA IACAACAG	8194
354	UGUUGUCC U CCAAUUUG	795	CAAAUUGG CUGAUGAG GCCGUUAGGC CGAA IGACAACA	8195
356	UUGUCCUC C AAUUUGUC	796	GACAAAUU CUGAUGAG GCCGUUAGGC CGAA IAGGACAA	8196
357	UGUCCUCC A AUUUGUCC	797	GGACAAAU CUGAUGAG GCCGUUAGGC CGAA IGAGGACA	8197
365	AAUUUGUC C UGGUUAUC	798	GAUAACCA CUGAUGAG GCCGUUAGGC CGAA IACAAAUU	8198

366	AUUUGUCC U GGUUAUCG	799	CGAUAACC CUGAUGAG GCCGUUAGGC CGAA IGACAAAU	8199
376	GUUAUCGC U GGAUGUGU	800	ACACAUCC CUGAUGAG GCCGUUAGGC CGAA ICGAUAAC	8200
386	GAUGUGUC U GCGGCGUU	801	AACGCCGC CUGAUGAG GCCGUUAGGC CGAA IACACAUC	8201
400	GUUUUAUC A UCUUCCUC	802	GAGGAAGA CUGAUGAG GCCGUUAGGC CGAA IAUAAAAC	8202
403	UUAUCAUC U UCCUCUGC	803	GCAGAGGA CUGAUGAG GCCGUUAGGC CGAA IAUGAUAA	8203
406	UCAUCUUC C UCUGCAUC	804	GAUGCAGA CUGAUGAG GCCGUUAGGC CGAA IAAGAUGA	8204
407	CAUCUUCC U CUGCAUCC	805	GGAUGCAG CUGAUGAG GCCGUUAGGC CGAA IGAAGAUG	8205
409	UCUUCCUC U GCAUCCUG	806	CAGGAUGC CUGAUGAG GCCGUUAGGC CGAA IAGGAAGA	8206
412	UCCUCUGC A UCCUGCUG	807	CAGCAGGA CUGAUGAG GCCGUUAGGC CGAA ICAGAGGA	8207
415	UCUGCAUC C UGCUGCUA	808	UAGCAGCA CUGAUGAG GCCGUUAGGC CGAA IAUGCAGA	8208
416	CUGCAUCC U GCUGCUAU	809	AUAGCAGC CUGAUGAG GCCGUUAGGC CGAA IGAUGCAG	8209
419	CAUCCUGC U GCUAUGCC	810	GGCAUAGC CUGAUGAG GCCGUUAGGC CGAA ICAGGAUG	8210
422	CCUGCUGC U AUGCCUCA	811	UGAGGCAU CUGAUGAG GCCGUUAGGC CGAA ICAGCAGG	8211
427	UGCUAUGC C UCAUCUUC	812	GAAGAUGA CUGAUGAG GCCGUUAGGC CGAA ICAUAGCA	8212
428	GCUAUGCC U CAUCUUCU	813	AGAAGAUG CUGAUGAG GCCGUUAGGC CGAA IGCAUAGC	8213
430	UAUGCCUC A UCUUCUUG	814	CAAGAAGA CUGAUGAG GCCGUUAGGC CGAA IAGGCAUA	8214
433	GCCUCAUC U UCUUGUUG	815	CAACAAGA CUGAUGAG GCCGUUAGGC CGAA IAUGAGGC	8215
436	UCAUCUUC U UGUUGGUU	816	AACCAACA CUGAUGAG GCCGUUAGGC CGAA IAAGAUGA	8216
446	GUUGGUUC U UCUGGACU	817	AGUCCAGA CUGAUGAG GCCGUUAGGC CGAA IAACCAAC	8217
449	GGUUCUUC U GGACUAUC	818	GAUAGUCC CUGAUGAG GCCGUUAGGC CGAA IAAGAACC	8218
454	UUCUGGAC U AUCAAGGU	819	ACCUUGAU CUGAUGAG GCCGUUAGGC CGAA IUCCAGAA	8219
458	GGACUAUC A AGGUAUGU	820	ACAUACCU CUGAUGAG GCCGUUAGGC CGAA IAUAGUCC	8220
470	UAUGUUGC C CGUUUGUC	821	GACAAACG CUGAUGAG GCCGUUAGGC CGAA ICAACAUA	8221
471	AUGUUGCC C GUUUGUCC	822	GGACAAAC CUGAUGAG GCCGUUAGGC CGAA IGCAACAU	8222
479	CGUUUGUC C UCUAAUUC	823	GAAUUAGA CUGAUGAG GCCGUUAGGC CGAA IACAAACG	8223
480	GUUUGUCC U CUAAUUCC	824	GGAAUUAG CUGAUGAG GCCGUUAGGC CGAA IGACAAAC	8224
482	UUGUCCUC U AAUUCCAG	825	CUGGAAUU CUGAUGAG GCCGUUAGGC CGAA IAGGACAA	8225
488	UCUAAUUC C AGGAUCAU	826	AUGAUCCU CUGAUGAG GCCGUUAGGC CGAA IAAUUAGA	8226
489	CUAAUUCC A GGAUCAUC	827	GAUGAUCC CUGAUGAG GCCGUUAGGC CGAA IGAAUUAG	8227
495	CCAGGAUC A UCAACAAC	828	GUUGUUGA CUGAUGAG GCCGUUAGGC CGAA IAUCCUGG	8228
498	GGAUCAUC A ACAACCAG	829	CUGGUUGU CUGAUGAG GCCGUUAGGC CGAA IAUGAUCC	8229
501	UCAUCAAC A ACCAGCAC	830	GUGCUGGU CUGAUGAG GCCGUUAGGC CGAA IUUGAUGA	8230
504	UCAACAAC C AGCACCGG	831	CCGGUGCU CUGAUGAG GCCGUUAGGC CGAA IUUGUUGA	8231
505	CAACAACC A GCACCGGA	832	UCCGGUGC CUGAUGAG GCCGUUAGGC CGAA IGUUGUUG	8232
508	CAACCAGC A CCGGACCA	833	UGGUCCGG CUGAUGAG GCCGUUAGGC CGAA ICUGGUUG	8233
510	ACCAGCAC C GGACCAUG	834	CAUGGUCC CUGAUGAG GCCGUUAGGC CGAA IUGCUGGU	8234
515	CACCGGAC C AUGCAAAA	835	UUUUGCAU CUGAUGAG GCCGUUAGGC CGAA IUCCGGUG	8235
516	ACCGGACC A UGCAAAAC	836	GUUUUGCA CUGAUGAG GCCGUUAGGC CGAA IGUCCGGU	8236
520	GACCAUGC A AAACCUGC	837	GCAGGUUU CUGAUGAG GCCGUUAGGC CGAA ICAUGGUC	8237
525	UGCAAAAC C UGCACAAC	838	GUUGUGCA CUGAUGAG GCCGUUAGGC CGAA IUUUUGCA	8238
526	GCAAAACC U GCACAACU	839	AGUUGUGC CUGAUGAG GCCGUUAGGC CGAA IGUUUUGC	8239
529	AAACCUGC A CAACUCCU	840	AGGAGUUG CUGAUGAG GCCGUUAGGC CGAA ICAGGUUU	8240
531	ACCUGCAC A ACUCCUGC	841	GCAGGAGU CUGAUGAG GCCGUUAGGC CGAA IUGCAGGU	8241
534	UGCACAAC U CCUGCUCA	842	UGAGCAGG CUGAUGAG GCCGUUAGGC CGAA IUUGUGCA	8242
536	CACAACUC C UGCUCAAG	843	CUUGAGCA CUGAUGAG GCCGUUAGGC CGAA IAGUUGUG	8243
537	ACAACUCC U GCUCAAGG	844	CCUUGAGC CUGAUGAG GCCGUUAGGC CGAA IGAGUUGU	8244
540	ACUCCUGC U CAAGGAAC	845	GUUCCUUG CUGAUGAG GCCGUUAGGC CGAA ICAGGAGU	8245
542	UCCUGCUC A AGGAACCU	846	AGGUUCCU CUGAUGAG GCCGUUAGGC CGAA IAGCAGGA	8246
549	CAAGGAAC C UCUAUGUU	847	AACAUAGA CUGAUGAG GCCGUUAGGC CGAA IUUCCUUG	8247
550	AAGGAACC U CUAUGUUU	848	AAACAUAG CUGAUGAG GCCGUUAGGC CGAA IGUUCCUU	8248
552	GGAACCUC U AUGUUUCC	849	GGAAACAU CUGAUGAG GCCGUUAGGC CGAA IAGGUUCC	8249

560	UAUGUUUC C CUCAUGUU	850	AACAUGAG CUGAUGAG GCCGUUAGGC CGAA IAAACAUA	8250
561	AUGUUUCC C UCAUGUUG	851	CAACAUGA CUGAUGAG GCCGUUAGGC CGAA IGAAACAU	8251
562	UGUUUCCC U CAUGUUGC	852	GCAACAUG CUGAUGAG GCCGUUAGGC CGAA IGGAAACA	8252
564	UUUCCCUC A UGUUGCUG	853	CAGCAACA CUGAUGAG GCCGUUAGGC CGAA IAGGGAAA	8253
571	CAUGUUGC U GUACAAAA	854	UUUUGUAC CUGAUGAG GCCGUUAGGC CGAA ICAACAUG	8254
576	UGCUGUAC A AAACCUAC	855	GUAGGUUU CUGAUGAG GCCGUUAGGC CGAA IUACAGCA	8255
581	UACAAAAC C UACGGACG	856	CGUCCGUA CUGAUGAG GCCGUUAGGC CGAA IUUUUGUA	8256
582	ACAAAACC U ACGGACGG	857	CCGUCCGU CUGAUGAG GCCGUUAGGC CGAA IGUUUUGU	8257
595	ACGGAAAC U GCACCUGU	858	ACAGGUGC CUGAUGAG GCCGUUAGGC CGAA IUUUCCGU	8258
598	GAAACUGC A CCUGUAUU	859	AAUACAGG CUGAUGAG GCCGUUAGGC CGAA ICAGUUUC	8259
600	AACUGCAC C UGUAUUCC	860	GGAAUACA CUGAUGAG GCCGUUAGGC CGAA IUGCAGUU	8260
601	ACUGCACC U GUAUUCCC	861	GGGAAUAC CUGAUGAG GCCGUUAGGC CGAA IGUGCAGU	8261
608	CUGUAUUC C CAUCCCAU	862	AUGGGAUG CUGAUGAG GCCGUUAGGC CGAA IAAUACAG	8262
609	UGUAUUCC C AUCCCAUC	863	GAUGGGAU CUGAUGAG GCCGUUAGGC CGAA IGAAUACA	8263
610	GUAUUCCC A UCCCAUCA	864	UGAUGGGA CUGAUGAG GCCGUUAGGC CGAA IGGAAUAC	8264
613	UUCCCAUC C CAUCAUCU	865	AGAUGAUG CUGAUGAG GCCGUUAGGC CGAA IAUGGGAA	8265
614	UCCCAUCC C AUCAUCUU	866	AAGAUGAU CUGAUGAG GCCGUUAGGC CGAA IGAUGGGA	8266
615	CCCAUCCC A UCAUCUUG	867	CAAGAUGA CUGAUGAG GCCGUUAGGC CGAA IGGAUGGG	8267
618	AUCCCAUC A UCUUGGGC	868	GCCCAAGA CUGAUGAG GCCGUUAGGC CGAA IAUGGGAU	8268
621	CCAUCAUC U UGGGCUUU	869	AAAGCCCA CUGAUGAG GCCGUUAGGC CGAA IAUGAUGG	8269
627	UCUUGGGC U UUCGCAAA	870	UUUGCGAA CUGAUGAG GCCGUUAGGC CGAA ICCCAAGA	8270
633	GCUUUCGC A AAAUACCU	871	AGGUAUUU CUGAUGAG GCCGUUAGGC CGAA ICGAAAGC	8271
640	CAAAAUAC C UAUGGGAG	872	CUCCCAUA CUGAUGAG GCCGUUAGGC CGAA IUAUUUUG	8272
641	AAAAUACC U AUGGGAGU	873	ACUCCCAU CUGAUGAG GCCGUUAGGC CGAA IGUAUUUU	8273
654	GAGUGGGC C UCAGUCCG	874	CGGACUGA CUGAUGAG GCCGUUAGGC CGAA ICCCACUC	8274
655	AGUGGGCC U CAGUCCGU	875	ACGGACUG CUGAUGAG GCCGUUAGGC CGAA IGCCCACU	8275
657	UGGGCCUC A GUCCGUUU	876	AAACGGAC CUGAUGAG GCCGUUAGGC CGAA IAGGCCCA	8276
661	CCUCAGUC C GUUUCUCU	877	AGAGAAAC CUGAUGAG GCCGUUAGGC CGAA IACUGAGG	8277
667	UCCGUUUC U CUUGGCUC	878	GAGCCAAG CUGAUGAG GCCGUUAGGC CGAA IAAACGGA	8278
669	CGUUUCUC U UGGCUCAG	879	CUGAGCCA CUGAUGAG GCCGUUAGGC CGAA IAGAAACG	8279
674	CUCUUGGC U CAGUUUAC	880	GUAAACUG CUGAUGAG GCCGUUAGGC CGAA ICCAAGAG	8280
676	CUUGGCUC A GUUUACUA	881	UAGUAAAC CUGAUGAG GCCGUUAGGC CGAA IAGCCAAG	8281
683	CAGUUUAC U AGUGCCAU	882	AUGGCACU CUGAUGAG GCCGUUAGGC CGAA IUAAACUG	8282
689	ACUAGUGC C AUUUGUUC	883	GAACAAAU CUGAUGAG GCCGUUAGGC CGAA ICACUAGU	8283
690	CUAGUGCC A UUUGUUCA	884	UGAACAAA CUGAUGAG GCCGUUAGGC CGAA IGCACUAG	8284
698	AUUUGUUC A GUGGUUCG	885	CGAACCAC CUGAUGAG GCCGUUAGGC CGAA IAACAAAU	8285
713	CGUAGGGC U UUCCCCCA	886	UGGGGGAA CUGAUGAG GCCGUUAGGC CGAA ICCCUACG	8286
717	GGGCUUUC C CCCACUGU	887	ACAGUGGG CUGAUGAG GCCGUUAGGC CGAA IAAAGCCC	8287
718	GGCUUUCC C CCACUGUC	888	GACAGUGG CUGAUGAG GCCGUUAGGC CGAA IGAAAGCC	8288
719	GCUUUCCC C CACUGUCU	889	AGACAGUG CUGAUGAG GCCGUUAGGC CGAA IGGAAAGC	8289
720	CUUUCCCC C ACUGUCUG	890	CAGACAGU CUGAUGAG GCCGUUAGGC CGAA IGGGAAAG	8290
721	UUUCCCCC A CUGUCUGG	891	CCAGACAG CUGAUGAG GCCGUUAGGC CGAA IGGGGAAA	8291
723	UCCCCCAC U GUCUGGCU	892	AGCCAGAC CUGAUGAG GCCGUUAGGC CGAA IUGGGGGA	8292
727	CCACUGUC U GGCUUUCA	893	UGAAAGCC CUGAUGAG GCCGUUAGGC CGAA IACAGUGG	8293
731	UGUCUGGC U UUCAGUUA	894	UAACUGAA CUGAUGAG GCCGUUAGGC CGAA ICCAGACA	8294
735	UGGCUUUC A GUUAUAUG	895	CAUAUAAC CUGAUGAG GCCGUUAGGC CGAA IAAAGCCA	8295
764	UUGGGGGC C AAGUCUGU	896	ACAGACUU CUGAUGAG GCCGUUAGGC CGAA ICCCCCAA	8296
765	UGGGGGCC A AGUCUGUA	897	UACAGACU CUGAUGAG GCCGUUAGGC CGAA IGCCCCCA	8297
770	GCCAAGUC U GUACAACA	898	UGUUGUAC CUGAUGAG GCCGUUAGGC CGAA IACUUGGC	8298
775	GUCUGUAC A ACAUCUUG	899	CAAGAUGU CUGAUGAG GCCGUUAGGC CGAA IUACAGAC	8299
778	UGUACAAC A UCUUGAGU	900	ACUCAAGA CUGAUGAG GCCGUUAGGC CGAA IUUGUACA	8300

781	ACAACAUC U UGAGUCCC	901	GGGACUCA CUGAUGAG GCCGUUAGGC CGAA IAUGUUGU	8301
788	CUUGAGUC C CUUUAUGC	902	GCAUAAAG CUGAUGAG GCCGUUAGGC CGAA IACUCAAG	8302
789	UUGAGUCC C UUUAUGCC	903	GGCAUAAA CUGAUGAG GCCGUUAGGC CGAA IGACUCAA	8303
790	UGAGUCCC U UUAUGCCG	904	CGGCAUAA CUGAUGAG GCCGUUAGGC CGAA IGGACUCA	8304
797	CUUUAUGC C GCUGUUAC	905	GUAACAGC CUGAUGAG GCCGUUAGGC CGAA ICAUAAAG	8305
800	UAUGCCGC U GUUACCAA	906	UUGGUAAC CUGAUGAG GCCGUUAGGC CGAA ICGGCAUA	8306
806	GCUGUUAC C AAUUUUCU	907	AGAAAAUU CUGAUGAG GCCGUUAGGC CGAA IUAACAGC	8307
807	CUGUUACC A AUUUUCUU	908	AAGAAAAU CUGAUGAG GCCGUUAGGC CGAA IGUAACAG	8308
814	CAAUUUUC U UUUGUCUU	909	AAGACAAA CUGAUGAG GCCGUUAGGC CGAA IAAAAUUG	8309
821	CUUUUGUC U UUGGGUAU	910	AUACCCAA CUGAUGAG GCCGUUAGGC CGAA IACAAAAG	8310
832	GGGUAUAC A UUUAAACC	911	GGUUUAAA CUGAUGAG GCCGUUAGGC CGAA IUAUACCC	8311
840	AUUUAAAC C CUCACAAA	912	UUUGUGAG CUGAUGAG GCCGUUAGGC CGAA IUUUAAAU	8312
841	UUUAAACC C UCACAAAA	913	UUUUGUGA CUGAUGAG GCCGUUAGGC CGAA IGUUUAAA	8313
842	UUAAACCC U CACAAAAC	914	GUUUUGUG CUGAUGAG GCCGUUAGGC CGAA IGGUUUAA	8314
844	AAACCCUC A CAAAACAA	915	UUGUUUUG CUGAUGAG GCCGUUAGGC CGAA IAGGGUUU	8315
846	ACCCUCAC A AAACAAAA	916	UUUUGUUU CUGAUGAG GCCGUUAGGC CGAA IUGAGGGU	8316
851	CACAAAAC A AAAAGAUG	917	CAUCUUUU CUGAUGAG GCCGUUAGGC CGAA IUUUUGUG	8317
869	GGAUAUUC C CUUAACUU	918	AAGUUAAG CUGAUGAG GCCGUUAGGC CGAA IAAUAUCC	8318
870	GAUAUUCC C UUAACUUC	919	GAAGUUAA CUGAUGAG GCCGUUAGGC CGAA IGAAUAUC	8319
871	AUAUUCCC U UAACUUCA	920	UGAAGUUA CUGAUGAG GCCGUUAGGC CGAA IGGAAUAU	8320
876	CCCUUAAC U UCAUGGGA	921	UCCCAUGA CUGAUGAG GCCGUUAGGC CGAA IUUAAGGG	8321
879	UUAACUUC A UGGGAUAU	922	AUAUCCCA CUGAUGAG GCCGUUAGGC CGAA IAAGUUAA	8322
906	GUUGGGGC A CAUUGCCA	923	UGGCAAUG CUGAUGAG GCCGUUAGGC CGAA ICCCCAAC	8323
908	UGGGGCAC A UUGCCACA	924	UGUGGCAA CUGAUGAG GCCGUUAGGC CGAA IUGCCCCA	8324
913	CACAUUGC C ACAGGAAC	925	GUUCCUGU CUGAUGAG GCCGUUAGGC CGAA ICAAUGUG	8325
914	ACAUUGCC A CAGGAACA	926	UGUUCCUG CUGAUGAG GCCGUUAGGC CGAA IGCAAUGU	8326
916	AUUGCCAC A GGAACAUA	927	UAUGUUCC CUGAUGAG GCCGUUAGGC CGAA IUGGCAAU	8327
922	ACAGGAAC A UAUUGUAC	928	GUACAAUA CUGAUGAG GCCGUUAGGC CGAA IUUCCUGU	8328
931	UAUUGUAC A AAAAAUCA	929	UGAUUUUU CUGAUGAG GCCGUUAGGC CGAA IUACAAUA	8329
939	AAAAAAUC A AAAUGUGU	930	ACACAUUU CUGAUGAG GCCGUUAGGC CGAA IAUUUUUU	8330
958	UAGGAAAC U UCCUGUAA	931	UUACAGGA CUGAUGAG GCCGUUAGGC CGAA IUUUCCUA	8331
961	GAAACUUC C UGUAAACA	932	UGUUUACA CUGAUGAG GCCGUUAGGC CGAA IAAGUUUC	8332
962	AAACUUCC U GUAAACAG	933	CUGUUUAC CUGAUGAG GCCGUUAGGC CGAA IGAAGUUU	8333
969	CUGUAAAC A GGCCUAUU	934	AAUAGGCC CUGAUGAG GCCGUUAGGC CGAA IUUUACAG	8334
973	AAACAGGC C UAUUGAUU	935	AAUCAAUA CUGAUGAG GCCGUUAGGC CGAA ICCUGUUU	8335
974	AACAGGCC U AUUGAUUG	936	CAAUCAAU CUGAUGAG GCCGUUAGGC CGAA IGCCUGUU	8336
994	AGUAUGUC A ACGAAUUG	937	CAAUUCGU CUGAUGAG GCCGUUAGGC CGAA IACAUACU	8337
1009	UGUGGGUC U UUUGGGGU	938	ACCCCAAA CUGAUGAG GCCGUUAGGC CGAA IACCCACA	8338
1022	GGGUUUGC C GCCCCUUU	939	AAAGGGGC CUGAUGAG GCCGUUAGGC CGAA ICAAACCC	8339
1025	UUUGCCGC C CCUUUCAC	940	GUGAAAGG CUGAUGAG GCCGUUAGGC CGAA ICGGCAAA	8340
1026	UUGCCGCC C CUUUCACG	941	CGUGAAAG CUGAUGAG GCCGUUAGGC CGAA IGCGGCAA	8341
1027	UGCCGCCC C UUUCACGC	942	GCGUGAAA CUGAUGAG GCCGUUAGGC CGAA IGGCGGCA	8342
1028	GCCGCCCC U UUCACGCA	943	UGCGUGAA CUGAUGAG GCCGUUAGGC CGAA IGGGCGGC	8343
1032	CCCCUUUC A CGCAAUGU	944	ACAUUGCG CUGAUGAG GCCGUUAGGC CGAA IAAAGGGG	8344
1036	UUUCACGC A AUGUGGAU	945	AUCCACAU CUGAUGAG GCCGUUAGGC CGAA ICGUGAAA	8345
1049	GGAUAUUC U GCUUUAAU	946	AUUAAAGC CUGAUGAG GCCGUUAGGC CGAA IAAUAUCC	8346
1052	UAUUCUGC U UUAAUGCC	947	GGCAUUAA CUGAUGAG GCCGUUAGGC CGAA ICAGAAUA	8347
1060	UUUAAUGC C UUUAUAUG	948	CAUAUAAA CUGAUGAG GCCGUUAGGC CGAA ICAUUAAA	8348
1061	UUAAUGCC U UUAUAUGC	949	GCAUAUAA CUGAUGAG GCCGUUAGGC CGAA IGCAUUAA	8349
1070	UUAUAUGC A UGCAUACA	950	UGUAUGCA CUGAUGAG GCCGUUAGGC CGAA ICAUAUAA	8350
1074	AUGCAUGC A UACAAGCA	951	UGCUUGUA CUGAUGAG GCCGUUAGGC CGAA ICAUGCAU	8351

1078	AUGCAUAC A AGCAAAAC	952	GUUUUGCU CUGAUGAG GCCGUUAGGC CGAA IUAUGCAU	8352
1082	AUACAAGC A AAACAGGC	953	GCCUGUUU CUGAUGAG GCCGUUAGGC CGAA ICUUGUAU	8353
1087	AGCAAAAC A GGCUUUUA	954	UAAAAGCC CUGAUGAG GCCGUUAGGC CGAA IUUUUGCU	8354
1091	AAACAGGC U UUUACUUU	955	AAAGUAAA CUGAUGAG GCCGUUAGGC CGAA ICCUGUUU	8355
1097	GCUUUUAC U UUCUCGCC	956	GGCGAGAA CUGAUGAG GCCGUUAGGC CGAA IUAAAAGC	8356
1101	UUACUUUC U CGCCAACU	957	AGUUGGCG CUGAUGAG GCCGUUAGGC CGAA IAAAGUAA	8357
1105	UUUCUCGC C AACUUACA	958	UGUAAGUU CUGAUGAG GCCGUUAGGC CGAA ICGAGAAA	8358
1106	UUCUCGCC A ACUUACAA	959	UUGUAAGU CUGAUGAG GCCGUUAGGC CGAA IGCGAGAA	8359
1109	UCGCCAAC U UACAAGGC	960	GCCUUGUA CUGAUGAG GCCGUUAGGC CGAA IUUGGCGA	8360
1113	CAACUUAC A AGGCCUUU	961	AAAGGCCU CUGAUGAG GCCGUUAGGC CGAA IUAAGUUG	8361
1118	UACAAGGC C UUUCUAAG	962	CUUAGAAA CUGAUGAG GCCGUUAGGC CGAA ICCUUGUA	8362
1119	ACAAGGCC U UUCUAAGU	963	ACUUAGAA CUGAUGAG GCCGUUAGGC CGAA IGCCUUGU	8363
1123	GGCCUUUC U AAGUAAAC	964	GUUUACUU CUGAUGAG GCCGUUAGGC CGAA IAAAGGCC	8364
1132	AAGUAAAC A GUAUGUGA	965	UCACAUAC CUGAUGAG GCCGUUAGGC CGAA IUUUACUU	8365
1143	AUGUGAAC C UUUACCCC	966	GGGGUAAA CUGAUGAG GCCGUUAGGC CGAA IUUCACAU	8366
1144	UGUGAACC U UUACCCCG	967	CGGGGUAA CUGAUGAG GCCGUUAGGC CGAA IGUUCACA	8367
1149	ACCUUUAC C CCGUUGCU	968	AGCAACGG CUGAUGAG GCCGUUAGGC CGAA IUAAAGGU	8368
1150	CCUUUACC C CGUUGCUC	969	GAGCAACG CUGAUGAG GCCGUUAGGC CGAA IGUAAAGG	8369
1151	CUUUACCC C GUUGCUCG	970	CGAGCAAC CUGAUGAG GCCGUUAGGC CGAA IGGUAAAG	8370
1157	CCCGUUGC U CGGCAACG	971	CGUUGCCG CUGAUGAG GCCGUUAGGC CGAA ICAACGGG	8371
1162	UGCUCGGC A ACGGCCUG	972	CAGGCCGU CUGAUGAG GCCGUUAGGC CGAA ICCGAGCA	8372
1168	GCAACGGC C UGGUCUAU	973	AUAGACCA CUGAUGAG GCCGUUAGGC CGAA ICCGUUGC	8373
1169	CAACGGCC U GGUCUAUG	974	CAUAGACC CUGAUGAG GCCGUUAGGC CGAA IGCCGUUG	8374
1174	GCCUGGUC U AUGCCAAG	975	CUUGGCAU CUGAUGAG GCCGUUAGGC CGAA IACCAGGC	8375
1179	GUCUAUGC C AAGUGUUU	976	AAACACUU CUGAUGAG GCCGUUAGGC CGAA ICAUAGAC	8376
1180	UCUAUGCC A AGUGUUUG	977	CAAACACU CUGAUGAG GCCGUUAGGC CGAA IGCAUAGA	8377
1190	GUGUUUGC U GACGCAAC	978	GUUGCGUC CUGAUGAG GCCGUUAGGC CGAA ICAAACAC	8378
1196	GCUGACGC A ACCCCCAC	979	GUGGGGGU CUGAUGAG GCCGUUAGGC CGAA ICGUCAGC	8379
1199	GACGCAAC C CCCACUGG	980	CCAGUGGG CUGAUGAG GCCGUUAGGC CGAA IUUGCGUC	8380
1200	ACGCAACC C CCACUGGU	981	ACCAGUGG CUGAUGAG GCCGUUAGGC CGAA IGUUGCGU	8381
1201	CGCAACCC C CACUGGUU	982	AACCAGUG CUGAUGAG GCCGUUAGGC CGAA IGGUUGCG	8382
1202	GCAACCCC C ACUGGUUG	983	CAACCAGU CUGAUGAG GCCGUUAGGC CGAA IGGGUUGC	8383
1203	CAACCCCC A CUGGUUGG	984	CCAACCAG CUGAUGAG GCCGUUAGGC CGAA IGGGGUUG	8384
1205	ACCCCCAC U GGUUGGGG	985	CCCCAACC CUGAUGAG GCCGUUAGGC CGAA IUGGGGGU	8385
1215	GUUGGGGC U UGGCCAUA	986	UAUGGCCA CUGAUGAG GCCGUUAGGC CGAA ICCCCAAC	8386
1220	GGCUUGGC C AUAGGCCA	987	UGGCCUAU CUGAUGAG GCCGUUAGGC CGAA ICCAAGCC	8387
1221	GCUUGGCC A UAGGCCAU	988	AUGGCCUA CUGAUGAG GCCGUUAGGC CGAA IGCCAAGC	8388
1227	CCAUAGGC C AUCAGCGC	989	GCGCUGAU CUGAUGAG GCCGUUAGGC CGAA ICCUAUGG	8389
1228	CAUAGGCC A UCAGCGCA	990	UGCGCUGA CUGAUGAG GCCGUUAGGC CGAA IGCCUAUG	8390
1231	AGGCCAUC A GCGCAUGC	991	GCAUGCGC CUGAUGAG GCCGUUAGGC CGAA IAUGGCCU	8391
1236	AUCAGCGC A UGCGUGGA	992	UCCACGCA CUGAUGAG GCCGUUAGGC CGAA ICGCUGAU	8392
1247	CGUGGAAC C UUUGUGUC	993	GACACAAA CUGAUGAG GCCGUUAGGC CGAA IUUCCACG	8393
1248	GUGGAACC U UUGUGUCU	994	AGACACAA CUGAUGAG GCCGUUAGGC CGAA IGUUCCAC	8394
1256	UUUGUGUC U CCUCUGCC	995	GGCAGAGG CUGAUGAG GCCGUUAGGC CGAA IACACAAA	8395
1258	UGUGUCUC C UCUGCCGA	996	UCGGCAGA CUGAUGAG GCCGUUAGGC CGAA IAGACACA	8396
1259	GUGUCUCC U CUGCCGAU	997	AUCGGCAG CUGAUGAG GCCGUUAGGC CGAA IGAGACAC	8397
1261	GUCUCCUC U GCCGAUCC	998	GGAUCGGC CUGAUGAG GCCGUUAGGC CGAA IAGGAGAC	8398
1264	UCCUCUGC C GAUCCAUA	999	UAUGGAUC CUGAUGAG GCCGUUAGGC CGAA ICAGAGGA	8399
1269	UGCCGAUC C AUACCGCG	1000	CGCGGUAU CUGAUGAG GCCGUUAGGC CGAA IAUCGGCA	8400
1270	GCCGAUCC A UACCGCGG	1001	CCGCGGUA CUGAUGAG GCCGUUAGGC CGAA IGAUCGGC	8401
1274	AUCCAUAC C GCGGAACU	1002	AGUUCCGC CUGAUGAG GCCGUUAGGC CGAA IUAUGGAU	8402

1282	CGCGGAAC U CCUAGCCG	1003	CGGCUAGG CUGAUGAG GCCGUUAGGC CGAA IUUCCGCG	8403
1284	CGGAACUC C UAGCCGCU	1004	AGCGGCUA CUGAUGAG GCCGUUAGGC CGAA IAGUUCCG	8404
1285	GGAACUCC U AGCCGCUU	1005	AAGCGGCU CUGAUGAG GCCGUUAGGC CGAA IGAGUUCC	8405
1289	CUCCUAGC C GCUUGUUU	1006	AAACAAGC CUGAUGAG GCCGUUAGGC CGAA ICUAGGAG	8406
1292	CUAGCCGC U UGUUUUGC	1007	GCAAAACA CUGAUGAG GCCGUUAGGC CGAA ICGGCUAG	8407
1301	UGUUUUGC U CGCAGCAG	1008	CUGCUGCG CUGAUGAG GCCGUUAGGC CGAA ICAAAACA	8408
1305	UUGCUCGC A GCAGGUCU	1009	AGACCUGC CUGAUGAG GCCGUUAGGC CGAA ICGAGCAA	8409
1308	CUCGCAGC A GGUCUGGG	1010	CCCAGACC CUGAUGAG GCCGUUAGGC CGAA ICUGCGAG	8410
1313	AGCAGGUC U GGGGCAAA	1011	UUUGCCCC CUGAUGAG GCCGUUAGGC CGAA IACCUGCU	8411
1319	UCUGGGGC A AAACUCAU	1012	AUGAGUUU CUGAUGAG GCCGUUAGGC CGAA ICCCCAGA	8412
1324	GGCAAAAC U CAUCGGGA	1013	UCCCGAUG CUGAUGAG GCCGUUAGGC CGAA IUUUUGCC	8413
1326	CAAAACUC A UCGGGACU	1014	AGUCCCGA CUGAUGAG GCCGUUAGGC CGAA IAGUUUUG	8414
1334	AUCGGGAC U GACAAUUC	1015	GAAUUGUC CUGAUGAG GCCGUUAGGC CGAA IUCCCGAU	8415
1338	GGACUGAC A AUUCUGUC	1016	GACAGAAU CUGAUGAG GCCGUUAGGC CGAA IUCAGUCC	8416
1343	GACAAUUC U GUCGUGCU	1017	AGCACGAC CUGAUGAG GCCGUUAGGC CGAA IAAUUGUC	8417
1351	UGUCGUGC U CUCCCGCA	1018	UGCGGGAG CUGAUGAG GCCGUUAGGC CGAA ICACGACA	8418
1353	UCGUGCUC U CCCGCAAA	1019	UUUGCGGG CUGAUGAG GCCGUUAGGC CGAA IAGCACGA	8419
1355	GUGCUCUC C CGCAAAUA	1020	UAUUUGCG CUGAUGAG GCCGUUAGGC CGAA IAGAGCAC	8420
1356	UGCUCUCC C GCAAAUAU	1021	AUAUUUGC CUGAUGAG GCCGUUAGGC CGAA IGAGAGCA	8421
1359	UCUCCCGC A AAUAUACA	1022	UGUAUAUU CUGAUGAG GCCGUUAGGC CGAA ICGGGAGA	8422
1367	AAAUAUAC A UCAUUUCC	1023	GGAAAUGA CUGAUGAG GCCGUUAGGC CGAA IUAUAUUU	8423
1370	UAUACAUC A UUUCCAUG	1024	CAUGGAAA CUGAUGAG GCCGUUAGGC CGAA IAUGUAUA	8424
1375	AUCAUUUC C AUGGCUGC	1025	GCAGCCAU CUGAUGAG GCCGUUAGGC CGAA IAAAUGAU	8425
1376	UCAUUUCC A UGGCUGCU	1026	AGCAGCCA CUGAUGAG GCCGUUAGGC CGAA IGAAAUGA	8426
1381	UCCAUGGC U GCUAGGCU	1027	AGCCUAGC CUGAUGAG GCCGUUAGGC CGAA ICCAUGGA	8427
1384	AUGGCUGC U AGGCUGUG	1028	CACAGCCU CUGAUGAG GCCGUUAGGC CGAA ICAGCCAU	8428
1389	UGCUAGGC U GUGCUGCC	1029	GGCAGCAC CUGAUGAG GCCGUUAGGC CGAA ICCUAGCA	8429
1394	GGCUGUGC U GCCAACUG	1030	CAGUUGGC CUGAUGAG GCCGUUAGGC CGAA ICACAGCC	8430
1397	UGUGCUGC C AACUGGAU	1031	AUCCAGUU CUGAUGAG GCCGUUAGGC CGAA ICAGCACA	8431
1398	GUGCUGCC A ACUGGAUC	1032	GAUCCAGU CUGAUGAG GCCGUUAGGC CGAA IGCAGCAC	8432
1401	CUGCCAAC U GGAUCCUA	1033	UAGGAUCC CUGAUGAG GCCGUUAGGC CGAA IUUGGCAG	8433
1407	ACUGGAUC C UACGCGGG	1034	CCCGCGUA CUGAUGAG GCCGUUAGGC CGAA IAUCCAGU	8434
1408	CUGGAUCC U ACGCGGGA	1035	UCCCGCGU CUGAUGAG GCCGUUAGGC CGAA IGAUCCAG	8435
1421	GGGACGUC C UUUGUUUA	1036	UAAACAAA CUGAUGAG GCCGUUAGGC CGAA IACGUCCC	8436
1422	GGACGUCC U UUGUUUAC	1037	GUAAACAA CUGAUGAG GCCGUUAGGC CGAA IGACGUCC	8437
1434	UUUACGUC C CGUCGGCG	1038	CGCCGACG CUGAUGAG GCCGUUAGGC CGAA IACGUAAA	8438
1435	UUACGUCC C GUCGGCGC	1039	GCGCCGAC CUGAUGAG GCCGUUAGGC CGAA IGACGUAA	8439
1444	GUCGGCGC U GAAUCCCG	1040	CGGGAUUC CUGAUGAG GCCGUUAGGC CGAA ICGCCGAC	8440
1450	GCUGAAUC C CGCGGACG	1041	CGUCCGCG CUGAUGAG GCCGUUAGGC CGAA IAUUCAGC	8441
1451	CUGAAUCC C GCGGACGA	1042	UCGUCCGC CUGAUGAG GCCGUUAGGC CGAA IGAUUCAG	8442
1461	CGGACGAC C CCUCCCGG	1043	CCGGGAGG CUGAUGAG GCCGUUAGGC CGAA IUCGUCCG	8443
1462	GGACGACC C CUCCCGGG	1044	CCCGGGAG CUGAUGAG GCCGUUAGGC CGAA IGUCGUCC	8444
1463	GACGACCC C UCCCGGGG	1045	CCCCGGGA CUGAUGAG GCCGUUAGGC CGAA IGGUCGUC	8445
1464	ACGACCCC U CCCGGGGC	1046	GCCCCGGG CUGAUGAG GCCGUUAGGC CGAA IGGGUCGU	8446
1466	GACCCCUC C CGGGGCCG	1047	CGGCCCCG CUGAUGAG GCCGUUAGGC CGAA IAGGGGUC	8447
1467	ACCCCUCC C GGGGCCGC	1048	GCGGCCCC CUGAUGAG GCCGUUAGGC CGAA IGAGGGGU	8448
1473	CCCGGGGC C GCUUGGGG	1049	CCCCAAGC CUGAUGAG GCCGUUAGGC CGAA ICCCCGGG	8449
1476	GGGGCCGC U UGGGGCUC	1050	GAGCCCCA CUGAUGAG GCCGUUAGGC CGAA ICGGCCCC	8450
1483	CUUGGGGC U CUACCGCC	1051	GGCGGUAG CUGAUGAG GCCGUUAGGC CGAA ICCCCAAG	8451
1485	UGGGGCUC U ACCGCCCG	1052	CGGGCGGU CUGAUGAG GCCGUUAGGC CGAA IAGCCCCA	8452
1488	GGCUCUAC C GCCCGCUU	1053	AAGCGGGC CUGAUGAG GCCGUUAGGC CGAA IUAGAGCC	8453

1491	UCUACCGC C CGCUUCUC	1054	GAGAAGCG CUGAUGAG GCCGUUAGGC CGAA ICGGUAGA	8454
1492	CUACCGCC C GCUUCUCC	1055	GGAGAAGC CUGAUGAG GCCGUUAGGC CGAA IGCGGUAG	8455
1495	CCGCCCGC U UCUCCGCC	1056	GGCGGAGA CUGAUGAG GCCGUUAGGC CGAA ICGGGCGG	8456
1498	CCCGCUUC U CCGCCUAU	1057	AUAGGCGG CUGAUGAG GCCGUUAGGC CGAA IAAGCGGG	8457
1500	CGCUUCUC C GCCUAUUG	1058	CAAUAGGC CUGAUGAG GCCGUUAGGC CGAA IAGAAGCG	8458
1503	UUCUCCGC C UAUUGUAC	1059	GUACAAUA CUGAUGAG GCCGUUAGGC CGAA ICGGAGAA	8459
1504	UCUCCGCC U AUUGUACC	1060	GGUACAAU CUGAUGAG GCCGUUAGGC CGAA IGCGGAGA	8460
1512	UAUUGUAC C GACCGUCC	1061	GGACGGUC CUGAUGAG GCCGUUAGGC CGAA IUACAAUA	8461
1516	GUACCGAC C GUCCACGG	1062	CCGUGGAC CUGAUGAG GCCGUUAGGC CGAA IUCGGUAC	8462
1520	CGACCGUC C ACGGGGCG	1063	CGCCCCGU CUGAUGAG GCCGUUAGGC CGAA IACGGUCG	8463
1521	GACCGUCC A CGGGGCGC	1064	GCGCCCCG CUGAUGAG GCCGUUAGGC CGAA IGACGGUC	8464
1530	CGGGGCGC A CCUCUCUU	1065	AAGAGAGG CUGAUGAG GCCGUUAGGC CGAA ICGCCCCG	8465
1532	GGGCGCAC C UCUCUUUA	1066	UAAAGAGA CUGAUGAG GCCGUUAGGC CGAA IUGCGCCC	8466
1533	GGCGCACC U CUCUUUAC	1067	GUAAAGAG CUGAUGAG GCCGUUAGGC CGAA IGUGCGCC	8467
1535	CGCACCUC U CUUUACGC	1068	GCGUAAAG CUGAUGAG GCCGUUAGGC CGAA IAGGUGCG	8468
1537	CACCUCUC U UUACGCGG	1069	CCGCGUAA CUGAUGAG GCCGUUAGGC CGAA IAGAGGUG	8469
1548	ACGCGGAC U CCCCGUCU	1070	AGACGGGG CUGAUGAG GCCGUUAGGC CGAA IUCCGCGU	8470
1550	GCGGACUC C CCGUCUGU	1071	ACAGACGG CUGAUGAG GCCGUUAGGC CGAA IAGUCCGC	8471
1551	CGGACUCC C CGUCUGUG	1072	CACAGACG CUGAUGAG GCCGUUAGGC CGAA IGAGUCCG	8472
1552	GGACUCCC C GUCUGUGC	1073	GCACAGAC CUGAUGAG GCCGUUAGGC CGAA IGGAGUCC	8473
1556	UCCCCGUC U GUGCCUUC	1074	GAAGGCAC CUGAUGAG GCCGUUAGGC CGAA IACGGGGA	8474
1561	GUCUGUGC C UUCUCAUC	1075	GAUGAGAA CUGAUGAG GCCGUUAGGC CGAA ICACAGAC	8475
1562	UCUGUGCC U UCUCAUCU	1076	AGAUGAGA CUGAUGAG GCCGUUAGGC CGAA IGCACAGA	8476
1565	GUGCCUUC U CAUCUGCC	1077	GGCAGAUG CUGAUGAG GCCGUUAGGC CGAA IAAGGCAC	8477
1567	GCCUUCUC A UCUGCCGG	1078	CCGGCAGA CUGAUGAG GCCGUUAGGC CGAA IAGAAGGC	8478
1570	UUCUCAUC U GCCGGACC	1079	GGUCCGGC CUGAUGAG GCCGUUAGGC CGAA IAUGAGAA	8479
1573	UCAUCUGC C GGACCGUG	1080	CACGGUCC CUGAUGAG GCCGUUAGGC CGAA ICAGAUGA	8480
1578	UGCCGGAC C GUGUGCAC	1081	GUGCACAC CUGAUGAG GCCGUUAGGC CGAA IUCCGGCA	8481
1585	CCGUGUGC A CUUCGCUU	1082	AAGCGAAG CUGAUGAG GCCGUUAGGC CGAA ICACACGG	8482
1587	GUGUGCAC U UCGCUUCA	1083	UGAAGCGA CUGAUGAG GCCGUUAGGC CGAA IUGCACAC	8483
1592	CACUUCGC U UCACCUCU	1084	AGAGGUGA CUGAUGAG GCCGUUAGGC CGAA ICGAAGUG	8484
1595	UUCGCUUC A CCUCUGCA	1085	UGCAGAGG CUGAUGAG GCCGUUAGGC CGAA IAAGCGAA	8485
1597	CGCUUCAC C UCUGCACG	1086	CGUGCAGA CUGAUGAG GCCGUUAGGC CGAA IUGAAGCG	8486
1598	GCUUCACC U CUGCACGU	1087	ACGUGCAG CUGAUGAG GCCGUUAGGC CGAA IGUGAAGC	8487
1600	UUCACCUC U GCACGUCG	1088	CGACGUGC CUGAUGAG GCCGUUAGGC CGAA IAGGUGAA	8488
1603	ACCUCUGC A CGUCGCAU	1089	AUGCGACG CUGAUGAG GCCGUUAGGC CGAA ICAGAGGU	8489
1610	CACGUCGC A UGGAGACC	1090	GGUCUCCA CUGAUGAG GCCGUUAGGC CGAA ICGACGUG	8490
1618	AUGGAGAC C ACCGUGAA	1091	UUCACGGU CUGAUGAG GCCGUUAGGC CGAA IUCUCCAU	8491
1619	UGGAGACC A CCGUGAAC	1092	GUUCACGG CUGAUGAG GCCGUUAGGC CGAA IGUCUCCA	8492
1621	GAGACCAC C GUGAACGC	1093	GCGUUCAC CUGAUGAG GCCGUUAGGC CGAA IUGGUCUC	8493
1630	GUGAACGC C CACAGGAA	1094	UUCCUGUG CUGAUGAG GCCGUUAGGC CGAA ICGUUCAC	8494
1631	UGAACGCC C ACAGGAAC	1095	GUUCCUGU CUGAUGAG GCCGUUAGGC CGAA IGCGUUCA	8495
1632	GAACGCCC A CAGGAACC	1096	GGUUCCUG CUGAUGAG GCCGUUAGGC CGAA IGGCGUUC	8496
1634	ACGCCCAC A GGAACCUG	1097	CAGGUUCC CUGAUGAG GCCGUUAGGC CGAA IUGGGCGU	8497
1640	ACAGGAAC C UGCCCAAG	1098	CUUGGGCA CUGAUGAG GCCGUUAGGC CGAA IUUCCUGU	8498
1641	CAGGAACC U GCCCAAGG	1099	CCUUGGGC CUGAUGAG GCCGUUAGGC CGAA IGUUCCUG	8499
1644	GAACCUGC C CAAGGUCU	1100	AGACCUUG CUGAUGAG GCCGUUAGGC CGAA ICAGGUUC	8500
1645	AACCUGCC C AAGGUCUU	1101	AAGACCUU CUGAUGAG GCCGUUAGGC CGAA IGCAGGUU	8501
1646	ACCUGCCC A AGGUCUUG	1102	CAAGACCU CUGAUGAG GCCGUUAGGC CGAA IGGCAGGU	8502
1652	CCAAGGUC U UGCAUAAG	1103	CUUAUGCA CUGAUGAG GCCGUUAGGC CGAA IACCUUGG	8503
1656	GGUCUUGC A UAAGAGGA	1104	UCCUCUUA CUGAUGAG GCCGUUAGGC CGAA ICAAGACC	8504

1666	AAGAGGAC U CUUGGACU	1105	AGUCCAAG CUGAUGAG GCCGUUAGGC CGAA IUCCUCUU	8505
1668	GAGGACUC U UGGACUUU	1106	AAAGUCCA CUGAUGAG GCCGUUAGGC CGAA IAGUCCUC	8506
1674	UCUUGGAC U UUCAGCAA	1107	UUGCUGAA CUGAUGAG GCCGUUAGGC CGAA IUCCAAGA	8507
1678	GGACUUUC A GCAAUGUC	1108	GACAUUGC CUGAUGAG GCCGUUAGGC CGAA IAAAGUCC	8508
1681	CUUUCAGC A AUGUCAAC	1109	GUUGACAU CUGAUGAG GCCGUUAGGC CGAA ICUGAAAG	8509
1687	GCAAUGUC A ACGACCGA	1110	UCGGUCGU CUGAUGAG GCCGUUAGGC CGAA IACAUUGC	8510
1693	UCAACGAC C GACCUUGA	1111	UCAAGGUC CUGAUGAG GCCGUUAGGC CGAA IUCGUUGA	8511
1697	CGACCGAC C UUGAGGCA	1112	UGCCUCAA CUGAUGAG GCCGUUAGGC CGAA IUCGGUCG	8512
1698	GACCGACC U UGAGGCAU	1113	AUGCCUCA CUGAUGAG GCCGUUAGGC CGAA IGUCGGUC	8513
1705	CUUGAGGC A UACUUCAA	1114	UUGAAGUA CUGAUGAG GCCGUUAGGC CGAA ICCUCAAG	8514
1709	AGGCAUAC U UCAAAGAC	1115	GUCUUUGA CUGAUGAG GCCGUUAGGC CGAA IUAUGCCU	8515
1712	CAUACUUC A AAGACUGU	1116	ACAGUCUU CUGAUGAG GCCGUUAGGC CGAA IAAGUAUG	8516
1718	UCAAAGAC U GUGUGUUU	1117	AAACACAC CUGAUGAG GCCGUUAGGC CGAA IUCUUUGA	8517
1769	UAAAGGUC U UUGUACUA	1118	UAGUACAA CUGAUGAG GCCGUUAGGC CGAA IACCUUUA	8518
1776	CUUUGUAC U AGGAGGCU	1119	AGCCUCCU CUGAUGAG GCCGUUAGGC CGAA IUACAAAG	8519
1784	UAGGAGGC U GUAGGCAU	1120	AUGCCUAC CUGAUGAG GCCGUUAGGC CGAA ICCUCCUA	8520
1791	CUGUAGGC A UAAAUUGG	1121	CCAAUUUA CUGAUGAG GCCGUUAGGC CGAA ICCUACAG	8521
1807	GUGUGUUC A CCAGCACC	1122	GGUGCUGG CUGAUGAG GCCGUUAGGC CGAA IAACACAC	8522
1809	GUGUUCAC C AGCACCAU	1123	AUGGUGCU CUGAUGAG GCCGUUAGGC CGAA IUGAACAC	8523
1810	UGUUCACC A GCACCAUG	1124	CAUGGUGC CUGAUGAG GCCGUUAGGC CGAA IGUGAACA	8524
1813	UCACCAGC A CCAUGCAA	1125	UUGCAUGG CUGAUGAG GCCGUUAGGC CGAA ICUGGUGA	8525
1815	ACCAGCAC C AUGCAACU	1126	AGUUGCAU CUGAUGAG GCCGUUAGGC CGAA IUGCUGGU	8526
1816	CCAGCACC A UGCAACUU	1127	AAGUUGCA CUGAUGAG GCCGUUAGGC CGAA IGUGCUGG	8527
1820	CACCAUGC A ACUUUUUC	1128	GAAAAAGU CUGAUGAG GCCGUUAGGC CGAA ICAUGGUG	8528
1823	CAUGCAAC U UUUUCACC	1129	GGUGAAAA CUGAUGAG GCCGUUAGGC CGAA IUUGCAUG	8529
1829	ACUUUUUC A CCUCUGCC	1130	GGCAGAGG CUGAUGAG GCCGUUAGGC CGAA IAAAAAGU	8530
1831	UUUUUCAC C UCUGCCUA	1131	UAGGCAGA CUGAUGAG GCCGUUAGGC CGAA IUGAAAAA	8531
1832	UUUUCACC U CUGCCUAA	1132	UUAGGCAG CUGAUGAG GCCGUUAGGC CGAA IGUGAAAA	8532
1834	UUCACCUC U GCCUAAUC	1133	GAUUAGGC CUGAUGAG GCCGUUAGGC CGAA IAGGUGAA	8533
1837	ACCUCUGC C UAAUCAUC	1134	GAUGAUUA CUGAUGAG GCCGUUAGGC CGAA ICAGAGGU	8534
1838	CCUCUGCC U AAUCAUCU	1135	AGAUGAUU CUGAUGAG GCCGUUAGGC CGAA IGCAGAGG	8535
1843	GCCUAAUC A UCUCAUGU	1136	ACAUGAGA CUGAUGAG GCCGUUAGGC CGAA IAUUAGGC	8536
1846	UAAUCAUC U CAUGUUCA	1137	UGAACAUG CUGAUGAG GCCGUUAGGC CGAA IAUGAUUA	8537
1848	AUCAUCUC A UGUUCAUG	1138	CAUGAACA CUGAUGAG GCCGUUAGGC CGAA IAGAUGAU	8538
1854	UCAUGUUC A UGUCCUAC	1139	GUAGGACA CUGAUGAG GCCGUUAGGC CGAA IAACAUGA	8539
1859	UUCAUGUC C UACUGUUC	1140	GAACAGUA CUGAUGAG GCCGUUAGGC CGAA IACAUGAA	8540
1860	UCAUGUCC U ACUGUUCA	1141	UGAACAGU CUGAUGAG GCCGUUAGGC CGAA IGACAUGA	8541
1863	UGUCCUAC U GUUCAAGC	1142	GCUUGAAC CUGAUGAG GCCGUUAGGC CGAA IUAGGACA	8542
1868	UACUGUUC A AGCCUCCA	1143	UGGAGGCU CUGAUGAG GCCGUUAGGC CGAA IAACAGUA	8543
1872	GUUCAAGC C UCCAAGCU	1144	AGCUUGGA CUGAUGAG GCCGUUAGGC CGAA ICUUGAAC	8544
1873	UUCAAGCC U CCAAGCUG	1145	CAGCUUGG CUGAUGAG GCCGUUAGGC CGAA IGCUUGAA	8545
1875	CAAGCCUC C AAGCUGUG	1146	CACAGCUU CUGAUGAG GCCGUUAGGC CGAA IAGGCUUG	8546
1876	AAGCCUCC A AGCUGUGC	1147	GCACAGCU CUGAUGAG GCCGUUAGGC CGAA IGAGGCUU	8547
1880	CUCCAAGC U GUGCCUUG	1148	CAAGGCAC CUGAUGAG GCCGUUAGGC CGAA ICUUGGAG	8548
1885	AGCUGUGC C UUGGGUGG	1149	CCACCCAA CUGAUGAG GCCGUUAGGC CGAA ICACAGCU	8549
1886	GCUGUGCC U UGGGUGGC	1150	GCCACCCA CUGAUGAG GCCGUUAGGC CGAA IGCACAGC	8550
1895	UGGGUGGC U UUGGGGCA	1151	UGCCCCAA CUGAUGAG GCCGUUAGGC CGAA ICCACCCA	8551
1903	UUUGGGGC A UGGACAUU	1152	AAUGUCCA CUGAUGAG GCCGUUAGGC CGAA ICCCCAAA	8552
1909	GCAUGGAC A UUGACCCG	1153	CGGGUCAA CUGAUGAG GCCGUUAGGC CGAA IUCCAUGC	8553
1915	ACAUUGAC C CGUAUAAA	1154	UUUAUACG CUGAUGAG GCCGUUAGGC CGAA IUCAAUGU	8554
1916	CAUUGACC C GUAUAAAG	1155	CUUUAUAC CUGAUGAG GCCGUUAGGC CGAA IGUCAAUG	8555

1935	UUUGGAGC U UCUGUGGA	1156	UCCACAGA CUGAUGAG GCCGUUAGGC CGAA ICUCCAAA	8556
1938	GGAGCUUC U GUGGAGUU	1157	AACUCCAC CUGAUGAG GCCGUUAGGC CGAA IAAGCUCC	8557
1949	GGAGUUAC U CUCUUUUU	1158	AAAAAGAG CUGAUGAG GCCGUUAGGC CGAA IUAACUCC	8558
1951	AGUUACUC U CUUUUUUG	1159	CAAAAAAG CUGAUGAG GCCGUUAGGC CGAA IAGUAACU	8559
1953	UUACUCUC U UUUUUGCC	1160	GGCAAAAA CUGAUGAG GCCGUUAGGC CGAA IAGAGUAA	8560
1961	UUUUUUGC C UUCUGACU	1161	AGUCAGAA CUGAUGAG GCCGUUAGGC CGAA ICAAAAAA	8561
1962	UUUUUGCC U UCUGACUU	1162	AAGUCAGA CUGAUGAG GCCGUUAGGC CGAA IGCAAAAA	8562
1965	UUGCCUUC U GACUUCUU	1163	AAGAAGUC CUGAUGAG GCCGUUAGGC CGAA IAAGGCAA	8563
1969	CUUCUGAC U UCUUUCCU	1164	AGGAAAGA CUGAUGAG GCCGUUAGGC CGAA IUCAGAAG	8564
1972	CUGACUUC U UUCCUUCU	1165	AGAAGGAA CUGAUGAG GCCGUUAGGC CGAA IAAGUCAG	8565
1976	CUUCUUUC C UUCUAUUC	1166	GAAUAGAA CUGAUGAG GCCGUUAGGC CGAA IAAAGAAG	8566
1977	UUCUUUCC U UCUAUUCG	1167	CGAAUAGA CUGAUGAG GCCGUUAGGC CGAA IGAAAGAA	8567
1980	UUUCCUUC U AUUCGAGA	1168	UCUCGAAU CUGAUGAG GCCGUUAGGC CGAA IAAGGAAA	8568
1991	UCGAGAUC U CCUCGACA	1169	UGUCGAGG CUGAUGAG GCCGUUAGGC CGAA IAUCUCGA	8569
1993	GAGAUCUC C UCGACACC	1170	GGUGUCGA CUGAUGAG GCCGUUAGGC CGAA IAGAUCUC	8570
1994	AGAUCUCC U CGACACCG	1171	CGGUGUCG CUGAUGAG GCCGUUAGGC CGAA IGAGAUCU	8571
1999	UCCUCGAC A CCGCCUCU	1172	AGAGGCGG CUGAUGAG GCCGUUAGGC CGAA IUCGAGGA	8572
2001	CUCGACAC C GCCUCUGC	1173	GCAGAGGC CUGAUGAG GCCGUUAGGC CGAA IUGUCGAG	8573
2004	GACACCGC C UCUGCUCU	1174	AGAGCAGA CUGAUGAG GCCGUUAGGC CGAA ICGGUGUC	8574
2005	ACACCGCC U CUGCUCUG	1175	CAGAGCAG CUGAUGAG GCCGUUAGGC CGAA IGCGGUGU	8575
2007	ACCGCCUC U GCUCUGUA	1176	UACAGAGC CUGAUGAG GCCGUUAGGC CGAA IAGGCGGU	8576
2010	GCCUCUGC U CUGUAUCG	1177	CGAUACAG CUGAUGAG GCCGUUAGGC CGAA ICAGAGGC	8577
2012	CUCUGCUC U GUAUCGGG	1178	CCCGAUAC CUGAUGAG GCCGUUAGGC CGAA IAGCAGAG	8578
2025	CGGGGGGC C UUAGAGUC	1179	GACUCUAA CUGAUGAG GCCGUUAGGC CGAA ICCCCCCG	8579
2026	GGGGGGCC U UAGAGUCU	1180	AGACUCUA CUGAUGAG GCCGUUAGGC CGAA IGCCCCCC	8580
2034	UUAGAGUC U CCGGAACA	1181	UGUUCCGG CUGAUGAG GCCGUUAGGC CGAA IACUCUAA	8581
2036	AGAGUCUC C GGAACAUU	1182	AAUGUUCC CUGAUGAG GCCGUUAGGC CGAA IAGACUCU	8582
2042	UCCGGAAC A UUGUUCAC	1183	GUGAACAA CUGAUGAG GCCGUUAGGC CGAA IUUCCGGA	8583
2049	CAUUGUUC A CCUCACCA	1184	UGGUGAGG CUGAUGAG GCCGUUAGGC CGAA IAACAAUG	8584
2051	UUGUUCAC C UCACCAUA	1185	UAUGGUGA CUGAUGAG GCCGUUAGGC CGAA IUGAACAA	8585
2052	UGUUCACC U CACCAUAC	1186	GUAUGGUG CUGAUGAG GCCGUUAGGC CGAA IGUGAACA	8586
2054	UUCACCUC A CCAUACGG	1187	CCGUAUGG CUGAUGAG GCCGUUAGGC CGAA IAGGUGAA	8587
2056	CACCUCAC C AUACGGCA	1188	UGCCGUAU CUGAUGAG GCCGUUAGGC CGAA IUGAGGUG	8588
2057	ACCUCACC A UACGGCAC	1189	GUGCCGUA CUGAUGAG GCCGUUAGGC CGAA IGUGAGGU	8589
2064	CAUACGGC A CUCAGGCA	1190	UGCCUGAG CUGAUGAG GCCGUUAGGC CGAA ICCGUAUG	8590
2066	UACGGCAC U CAGGCAAG	1191	CUUGCCUG CUGAUGAG GCCGUUAGGC CGAA IUGCCGUA	8591
2068	CGGCACUC A GGCAAGCU	1192	AGCUUGCC CUGAUGAG GCCGUUAGGC CGAA IAGUGCCG	8592
2072	ACUCAGGC A AGCUAUUC	1193	GAAUAGCU CUGAUGAG GCCGUUAGGC CGAA ICCUGAGU	8593
2076	AGGCAAGC U AUUCUGUG	1194	CACAGAAU CUGAUGAG GCCGUUAGGC CGAA ICUUGCCU	8594
2081	AGCUAUUC U GUGUUGGG	1195	CCCAACAC CUGAUGAG GCCGUUAGGC CGAA IAAUAGCU	8595
2105	GAUGAAUC U AGCCACCU	1196	AGGUGGCU CUGAUGAG GCCGUUAGGC CGAA IAUUCAUC	8596
2109	AAUCUAGC C ACCUGGGU	1197	ACCCAGGU CUGAUGAG GCCGUUAGGC CGAA ICUAGAUU	8597
2110	AUCUAGCC A CCUGGGUG	1198	CACCCAGG CUGAUGAG GCCGUUAGGC CGAA IGCUAGAU	8598
2112	CUAGCCAC C UGGGUGGG	1199	CCCACCCA CUGAUGAG GCCGUUAGGC CGAA IUGGCUAG	8599
2113	UAGCCACC U GGGUGGGA	1200	UCCCACCC CUGAUGAG GCCGUUAGGC CGAA IGUGGCUA	8600
2138	GGAAGAUC C AGCAUCCA	1201	UGGAUGCU CUGAUGAG GCCGUUAGGC CGAA IAUCUUCC	8601
2139	GAAGAUCC A GCAUCCAG	1202	CUGGAUGC CUGAUGAG GCCGUUAGGC CGAA IGAUCUUC	8602
2142	GAUCCAGC A UCCAGGGA	1203	UCCCUGGA CUGAUGAG GCCGUUAGGC CGAA ICUGGAUC	8603
2145	CCAGCAUC C AGGGAAUU	1204	AAUUCCCU CUGAUGAG GCCGUUAGGC CGAA IAUGCUGG	8604
2146	CAGCAUCC A GGGAAUUA	1205	UAAUUCCC CUGAUGAG GCCGUUAGGC CGAA IGAUGCUG	8605
2161	UAGUAGUC A GCUAUGUC	1206	GACAUAGC CUGAUGAG GCCGUUAGGC CGAA IACUACUA	8606

2164	UAGUCAGC U AUGUCAAC	1207	GUUGACAU CUGAUGAG GCCGUUAGGC CGAA ICUGACUA	8607
2170	GCUAUGUC A ACGUUAAU	1208	AUUAACGU CUGAUGAG GCCGUUAGGC CGAA IACAUAGC	8608
2185	AUAUGGGC C UAAAAAUC	1209	GAUUUUUA CUGAUGAG GCCGUUAGGC CGAA ICCCAUAU	8609
2186	UAUGGGCC U AAAAAUCA	1210	UGAUUUUU CUGAUGAG GCCGUUAGGC CGAA IGCCCAUA	8610
2194	UAAAAAUC A GACAACUA	1211	UAGUUGUC CUGAUGAG GCCGUUAGGC CGAA IAUUUUUA	8611
2198	AAUCAGAC A ACUAUUGU	1212	ACAAUAGU CUGAUGAG GCCGUUAGGC CGAA IUCUGAUU	8612
2201	CAGACAAC U AUUGUGGU	1213	ACCACAAU CUGAUGAG GCCGUUAGGC CGAA IUUGUCUG	8613
2213	GUGGUUUC A CAUUUCCU	1214	AGGAAAUG CUGAUGAG GCCGUUAGGC CGAA IAAACCAC	8614
2215	GGUUUCAC A UUUCCUGU	1215	ACAGGAAA CUGAUGAG GCCGUUAGGC CGAA IUGAAACC	8615
2220	CACAUUUC C UGUCUUAC	1216	GUAAGACA CUGAUGAG GCCGUUAGGC CGAA IAAAUGUG	8616
2221	ACAUUUCC U GUCUUACU	1217	AGUAAGAC CUGAUGAG GCCGUUAGGC CGAA IGAAAUGU	8617
2225	UUCCUGUC U UACUUUUG	1218	CAAAAGUA CUGAUGAG GCCGUUAGGC CGAA IACAGGAA	8618
2229	UGUCUUAC U UUUGGGCG	1219	CGCCCAAA CUGAUGAG GCCGUUAGGC CGAA IUAAGACA	8619
2244	CGAGAAAC U GUUCUUGA	1220	UCAAGAAC CUGAUGAG GCCGUUAGGC CGAA IUUUCUCG	8620
2249	AACUGUUC U UGAAUAUU	1221	AAUAUUCA CUGAUGAG GCCGUUAGGC CGAA IAACAGUU	8621
2265	UUGGUGUC U UUUGGAGU	1222	ACUCCAAA CUGAUGAG GCCGUUAGGC CGAA IACACCAA	8622
2284	GGAUUCGC A CUCCUCCU	1223	AGGAGGAG CUGAUGAG GCCGUUAGGC CGAA ICGAAUCC	8623
2286	AUUCGCAC U CCUCCUGC	1224	GCAGGAGG CUGAUGAG GCCGUUAGGC CGAA IUGCGAAU	8624
2288	UCGCACUC C UCCUGCAU	1225	AUGCAGGA CUGAUGAG GCCGUUAGGC CGAA IAGUGCGA	8625
2289	CGCACUCC U CCUGCAUA	1226	UAUGCAGG CUGAUGAG GCCGUUAGGC CGAA IGAGUGCG	8626
2291	CACUCCUC C UGCAUAUA	1227	UAUAUGCA CUGAUGAG GCCGUUAGGC CGAA IAGGAGUG	8627
2292	ACUCCUCC U GCAUAUAG	1228	CUAUAUGC CUGAUGAG GCCGUUAGGC CGAA IGAGGAGU	8628
2295	CCUCCUGC A UAUAGACC	1229	GGUCUAUA CUGAUGAG GCCGUUAGGC CGAA ICAGGAGG	8629
2303	AUAUAGAC C ACCAAAUG	1230	CAUUUGGU CUGAUGAG GCCGUUAGGC CGAA IUCUAUAU	8630
2304	UAUAGACC A CCAAAUGC	1231	GCAUUUGG CUGAUGAG GCCGUUAGGC CGAA IGUCUAUA	8631
2306	UAGACCAC C AAAUGCCC	1232	GGGCAUUU CUGAUGAG GCCGUUAGGC CGAA IUGGUCUA	8632
2307	AGACCACC A AAUGCCCC	1233	GGGGCAUU CUGAUGAG GCCGUUAGGC CGAA IGUGGUCU	8633
2313	CCAAAUGC C CCUAUCUU	1234	AAGAUAGG CUGAUGAG GCCGUUAGGC CGAA ICAUUUGG	8634
2314	CAAAUGCC C CUAUCUUA	1235	UAAGAUAG CUGAUGAG GCCGUUAGGC CGAA IGCAUUUG	8635
2315	AAAUGCCC C UAUCUUAU	1236	AUAAGAUA CUGAUGAG GCCGUUAGGC CGAA IGGCAUUU	8636
2316	AAUGCCCC U AUCUUAUC	1237	GAUAAGAU CUGAUGAG GCCGUUAGGC CGAA IGGGCAUU	8637
2320	CCCCUAUC U UAUCAACA	1238	UGUUGAUA CUGAUGAG GCCGUUAGGC CGAA IAUAGGGG	8638
2325	AUCUUAUC A ACACUUCC	1239	GGAAGUGU CUGAUGAG GCCGUUAGGC CGAA IAUAAGAU	8639
2328	UUAUCAAC A CUUCCGGA	1240	UCCGGAAG CUGAUGAG GCCGUUAGGC CGAA IUUGAUAA	8640
2330	AUCAACAC U UCCGGAAA	1241	UUUCCGGA CUGAUGAG GCCGUUAGGC CGAA IUGUUGAU	8641
2333	AACACUUC C GGAAACUA	1242	UAGUUUCC CUGAUGAG GCCGUUAGGC CGAA IAAGUGUU	8642
2340	CCGGAAAC U ACUGUUGU	1243	ACAACAGU CUGAUGAG GCCGUUAGGC CGAA IUUUCCGG	8643
2343	GAAACUAC U GUUGUUAG	1244	CUAACAAC CUGAUGAG GCCGUUAGGC CGAA IUAGUUUC	8644
2362	GAAGAGGC A GGUCCCCU	1245	AGGGGACC CUGAUGAG GCCGUUAGGC CGAA ICCUCUUC	8645
2367	GGCAGGUC C CCUAGAAG	1246	CUUCUAGG CUGAUGAG GCCGUUAGGC CGAA IACCUGCC	8646
2368	GCAGGUCC C CUAGAAGA	1247	UCUUCUAG CUGAUGAG GCCGUUAGGC CGAA IGACCUGC	8647
2369	CAGGUCCC C UAGAAGAA	1248	UUCUUCUA CUGAUGAG GCCGUUAGGC CGAA IGGACCUG	8648
2370	AGGUCCCC U AGAAGAAG	1249	CUUCUUCU CUGAUGAG GCCGUUAGGC CGAA IGGGACCU	8649
2382	AGAAGAAC U CCCUCGCC	1250	GGCGAGGG CUGAUGAG GCCGUUAGGC CGAA IUUCUUCU	8650
2384	AAGAACUC C CUCGCCUC	1251	GAGGCGAG CUGAUGAG GCCGUUAGGC CGAA IAGUUCUU	8651
2385	AGAACUCC C UCGCCUCG	1252	CGAGGCGA CUGAUGAG GCCGUUAGGC CGAA IGAGUUCU	8652
2386	GAACUCCC U CGCCUCGC	1253	GCGAGGCG CUGAUGAG GCCGUUAGGC CGAA IGGAGUUC	8653
2390	UCCCUCGC C UCGCAGAC	1254	GUCUGCGA CUGAUGAG GCCGUUAGGC CGAA ICGAGGGA	8654
2391	CCCUCGCC U CGCAGACG	1255	CGUCUGCG CUGAUGAG GCCGUUAGGC CGAA IGCGAGGG	8655
2395	CGCCUCGC A GACGAAGG	1256	CCUUCGUC CUGAUGAG GCCGUUAGGC CGAA ICGAGGCG	8656
2406	CGAAGGUC U CAAUCGCC	1257	GGCGAUUG CUGAUGAG GCCGUUAGGC CGAA IACCUUCG	8657

2408	AAGGUCUC A AUCGCCGC	1258	GCGGCGAU CUGAUGAG GCCGUUAGGC CGAA IAGACCUU	8658
2414	UCAAUCGC C GCGUCGCA	1259	UGCGACGC CUGAUGAG GCCGUUAGGC CGAA ICGAUUGA	8659
2422	CGCGUCGC A GAAGAUCU	1260	AGAUCUUC CUGAUGAG GCCGUUAGGC CGAA ICGACGCG	8660
2430	AGAAGAUC U CAAUCUCG	1261	CGAGAUUG CUGAUGAG GCCGUUAGGC CGAA IAUCUUCU	8661
2432	AAGAUCUC A AUCUCGGG	1262	CCCGAGAU CUGAUGAG GCCGUUAGGC CGAA IAGAUCUU	8662
2436	UCUCAAUC U CGGGAAUC	1263	GAUUCCCG CUGAUGAG GCCGUUAGGC CGAA IAUUGAGA	8663
2445	CGGGAAUC U CAAUGUUA	1264	UAACAUUG CUGAUGAG GCCGUUAGGC CGAA IAUUCCCG	8664
2447	GGAAUCUC A AUGUUAGU	1265	ACUAACAU CUGAUGAG GCCGUUAGGC CGAA IAGAUUCC	8665
2460	UAGUAUUC C UUGGACAC	1266	GUGUCCAA CUGAUGAG GCCGUUAGGC CGAA IAAUACUA	8666
2461	AGUAUUCC U UGGACACA	1267	UGUGUCCA CUGAUGAG GCCGUUAGGC CGAA IGAAUACU	8667
2467	CCUUGGAC A CAUAAGGU	1268	ACCUUAUG CUGAUGAG GCCGUUAGGC CGAA IUCCAAGG	8668
2469	UUGGACAC A UAAGGUGG	1269	CCACCUUA CUGAUGAG GCCGUUAGGC CGAA IUGUCCAA	8669
2483	UGGGAAAC U UUACGGGG	1270	CCCCGUAA CUGAUGAG GCCGUUAGGC CGAA IUUUCCCA	8670
2493	UACGGGGC U UUAUUCUU	1271	AAGAAUAA CUGAUGAG GCCGUUAGGC CGAA ICCCCGUA	8671
2500	CUUUAUUC U UCUACGGU	1272	ACCGUAGA CUGAUGAG GCCGUUAGGC CGAA IAAUAAAG	8672
2503	UAUUCUUC U ACGGUACC	1273	GGUACCGU CUGAUGAG GCCGUUAGGC CGAA IAAGAAUA	8673
2511	UACGGUAC C UUGCUUUA	1274	UAAAGCAA CUGAUGAG GCCGUUAGGC CGAA IUACCGUA	8674
2512	ACGGUACC U UGCUUUAA	1275	UUAAAGCA CUGAUGAG GCCGUUAGGC CGAA IGUACCGU	8675
2516	UACCUUGC U UUAAUCCU	1276	AGGAUUAA CUGAUGAG GCCGUUAGGC CGAA ICAAGGUA	8676
2523	CUUUAAUC C UAAAUGGC	1277	GCCAUUUA CUGAUGAG GCCGUUAGGC CGAA IAUUAAAG	8677
2524	UUUAAUCC U AAAUGGCA	1278	UGCCAUUU CUGAUGAG GCCGUUAGGC CGAA IGAUUAAA	8678
2532	UAAAUGGC A AACUCCUU	1279	AAGGAGUU CUGAUGAG GCCGUUAGGC CGAA ICCAUUUA	8679
2536	UGGCAAAC U CCUUCUUU	1280	AAAGAAGG CUGAUGAG GCCGUUAGGC CGAA IUUUGCCA	8680
2538	GCAAACUC C UUCUUUUC	1281	GAAAAGAA CUGAUGAG GCCGUUAGGC CGAA IAGUUUGC	8681
2539	CAAACUCC U UCUUUUCC	1282	GGAAAAGA CUGAUGAG GCCGUUAGGC CGAA IGAGUUUG	8682
2542	ACUCCUUC U UUUCCUGA	1283	UCAGGAAA CUGAUGAG GCCGUUAGGC CGAA IAAGGAGU	8683
2547	UUCUUUUC C UGACAUUC	1284	GAAUGUCA CUGAUGAG GCCGUUAGGC CGAA IAAAAGAA	8684
2548	UCUUUUCC U GACAUUCA	1285	UGAAUGUC CUGAUGAG GCCGUUAGGC CGAA IGAAAAGA	8685
2552	UUCCUGAC A UUCAUUUG	1286	CAAAUGAA CUGAUGAG GCCGUUAGGC CGAA IUCAGGAA	8686
2556	UGACAUUC A UUUGCAGG	1287	CCUGCAAA CUGAUGAG GCCGUUAGGC CGAA IAAUGUCA	8687
2562	UCAUUUGC A GGAGGACA	1288	UGUCCUCC CUGAUGAG GCCGUUAGGC CGAA ICAAAUGA	8688
2570	AGGAGGAC A UUGUUGAU	1289	AUCAACAA CUGAUGAG GCCGUUAGGC CGAA IUCCUCCU	8689
2589	AUGUAAGC A AUUUGUGG	1290	CCACAAAU CUGAUGAG GCCGUUAGGC CGAA ICUUACAU	8690
2601	UGUGGGGC C CCUUACAG	1291	CUGUAAGG CUGAUGAG GCCGUUAGGC CGAA ICCCCACA	8691
2602	GUGGGGCC C CUUACAGU	1292	ACUGUAAG CUGAUGAG GCCGUUAGGC CGAA IGCCCCAC	8692
2603	UGGGGCCC C UUACAGUA	1293	UACUGUAA CUGAUGAG GCCGUUAGGC CGAA IGGCCCCA	8693
2604	GGGGCCCC U UACAGUAA	1294	UUACUGUA CUGAUGAG GCCGUUAGGC CGAA IGGGCCCC	8694
2608	CCCCUUAC A GUAAAUGA	1295	UCAUUUAC CUGAUGAG GCCGUUAGGC CGAA IUAAGGGG	8695
2621	AUGAAAAC A GGAGACUU	1296	AAGUCUCC CUGAUGAG GCCGUUAGGC CGAA IUUUUCAU	8696
2628	CAGGAGAC U UAAAUUAA	1297	UUAAUUUA CUGAUGAG GCCGUUAGGC CGAA IUCUCCUG	8697
2638	AAAUUAAC U AUGCCUGC	1298	GCAGGCAU CUGAUGAG GCCGUUAGGC CGAA IUUAAUUU	8698
2643	AACUAUGC C UGCUAGGU	1299	ACCUAGCA CUGAUGAG GCCGUUAGGC CGAA ICAUAGUU	8699
2644	ACUAUGCC U GCUAGGUU	1300	AACCUAGC CUGAUGAG GCCGUUAGGC CGAA IGCAUAGU	8700
2647	AUGCCUGC U AGGUUUUA	1301	UAAAACCU CUGAUGAG GCCGUUAGGC CGAA ICAGGCAU	8701
2658	GUUUUAUC C CAAUGUUA	1302	UAACAUUG CUGAUGAG GCCGUUAGGC CGAA IAUAAAAC	8702
2659	UUUUAUCC C AAUGUUAC	1303	GUAACAUU CUGAUGAG GCCGUUAGGC CGAA IGAUAAAA	8703
2660	UUUAUCCC A AUGUUACU	1304	AGUAACAU CUGAUGAG GCCGUUAGGC CGAA IGGAUAAA	8704
2668	AAUGUUAC U AAAUAUUU	1305	AAAUAUUU CUGAUGAG GCCGUUAGGC CGAA IUAACAUU	8705
2679	AUAUUUGC C CUUAGAUA	1306	UAUCUAAG CUGAUGAG GCCGUUAGGC CGAA ICAAAUAU	8706
2680	UAUUUGCC C UUAGAUAA	1307	UUAUCUAA CUGAUGAG GCCGUUAGGC CGAA IGCAAAUA	8707
2681	AUUUGCCC U UAGAUAAA	1308	UUUAUCUA CUGAUGAG GCCGUUAGGC CGAA IGGCAAAU	8708

2696	AAGGGAUC A AACCGUAU	1309	AUACGGUU CUGAUGAG GCCGUUAGGC CGAA IAUCCCUU	8709
2700	GAUCAAAC C GUAUUAUC	1310	GAUAAUAC CUGAUGAG GCCGUUAGGC CGAA IUUUGAUC	8710
2709	GUAUUAUC C AGAGUAUG	1311	CAUACUCU CUGAUGAG GCCGUUAGGC CGAA IAUAAUAC	8711
2710	UAUUAUCC A GAGUAUGU	1312	ACAUACUC CUGAUGAG GCCGUUAGGC CGAA IGAUAAUA	8712
2727	AGUUAAUC A UUACUUCC	1313	GGAAGUAA CUGAUGAG GCCGUUAGGC CGAA IAUUAACU	8713
2732	AUCAUUAC U UCCAGACG	1314	CGUCUGGA CUGAUGAG GCCGUUAGGC CGAA IUAAUGAU	8714
2735	AUUACUUC C AGACGCGA	1315	UCGCGUCU CUGAUGAG GCCGUUAGGC CGAA IAAGUAAU	8715
2736	UUACUUCC A GACGCGAC	1316	GUCGCGUC CUGAUGAG GCCGUUAGGC CGAA IGAAGUAA	8716
2745	GACGCGAC A UUAUUUAC	1317	GUAAAUAA CUGAUGAG GCCGUUAGGC CGAA IUCGCGUC	8717
2754	UUAUUUAC A CACUCUUU	1318	AAAGAGUG CUGAUGAG GCCGUUAGGC CGAA IUAAAUAA	8718
2756	AUUUACAC A CUCUUUGG	1319	CCAAAGAG CUGAUGAG GCCGUUAGGC CGAA IUGUAAAU	8719
2758	UUACACAC U CUUUGGAA	1320	UUCCAAAG CUGAUGAG GCCGUUAGGC CGAA IUGUGUAA	8720
2760	ACACACUC U UUGGAAGG	1321	CCUUCCAA CUGAUGAG GCCGUUAGGC CGAA IAGUGUGU	8721
2777	CGGGGAUC U UAUAUAAA	1322	UUUAUAUA CUGAUGAG GCCGUUAGGC CGAA IAUCCCCG	8722
2794	AGAGAGUC C ACACGUAG	1323	CUACGUGU CUGAUGAG GCCGUUAGGC CGAA IACUCUCU	8723
2795	GAGAGUCC A CACGUAGC	1324	GCUACGUG CUGAUGAG GCCGUUAGGC CGAA IGACUCUC	8724
2797	GAGUCCAC A CGUAGCGC	1325	GCGCUACG CUGAUGAG GCCGUUAGGC CGAA IUGGACUC	8725
2806	CGUAGCGC C UCAUUUUG	1326	CAAAAUGA CUGAUGAG GCCGUUAGGC CGAA ICGCUACG	8726
2807	GUAGCGCC U CAUUUUGC	1327	GCAAAAUG CUGAUGAG GCCGUUAGGC CGAA IGCGCUAC	8727
2809	AGCGCCUC A UUUUGCGG	1328	CCGCAAAA CUGAUGAG GCCGUUAGGC CGAA IAGGCGCU	8728
2821	UGCGGGUC A CCAUAUUC	1329	GAAUAUGG CUGAUGAG GCCGUUAGGC CGAA IACCCGCA	8729
2823	CGGGUCAC C AUAUUCUU	1330	AAGAAUAU CUGAUGAG GCCGUUAGGC CGAA IUGACCCG	8730
2824	GGGUCACC A UAUUCUUG	1331	CAAGAAUA CUGAUGAG GCCGUUAGGC CGAA IGUGACCC	8731
2830	CCAUAUUC U UGGGAACA	1332	UGUUCCCA CUGAUGAG GCCGUUAGGC CGAA IAAUAUGG	8732
2838	UUGGGAAC A AGAUCUAC	1333	GUAGAUCU CUGAUGAG GCCGUUAGGC CGAA IUUCCCAA	8733
2844	ACAAGAUC U ACAGCAUG	1334	CAUGCUGU CUGAUGAG GCCGUUAGGC CGAA IAUCUUGU	8734
2847	AGAUCUAC A GCAUGGGA	1335	UCCCAUGC CUGAUGAG GCCGUUAGGC CGAA IUAGAUCU	8735
2850	UCUACAGC A UGGGAGGU	1336	ACCUCCCA CUGAUGAG GCCGUUAGGC CGAA ICUGUAGA	8736
2864	GGUUGGUC U UCCAAACC	1337	GGUUUGGA CUGAUGAG GCCGUUAGGC CGAA IACCAACC	8737
2867	UGGUCUUC C AAACCUCG	1338	CGAGGUUU CUGAUGAG GCCGUUAGGC CGAA IAAGACCA	8738
2868	GGUCUUCC A AACCUCGA	1339	UCGAGGUU CUGAUGAG GCCGUUAGGC CGAA IGAAGACC	8739
2872	UUCCAAAC C UCGAAAAG	1340	CUUUUCGA CUGAUGAG GCCGUUAGGC CGAA IUUUGGAA	8740
2873	UCCAAACC U CGAAAAGG	1341	CCUUUUCG CUGAUGAG GCCGUUAGGC CGAA IGUUUGGA	8741
2883	GAAAAGGC A UGGGGACA	1342	UGUCCCCA CUGAUGAG GCCGUUAGGC CGAA ICCUUUUC	8742
2891	AUGGGGAC A AAUCUUUC	1343	GAAAGAUU CUGAUGAG GCCGUUAGGC CGAA IUCCCCAU	8743
2896	GACAAAUC U UUCUGUCC	1344	GGACAGAA CUGAUGAG GCCGUUAGGC CGAA IAUUUGUC	8744
2900	AAUCUUUC U GUCCCCAA	1345	UUGGGGAC CUGAUGAG GCCGUUAGGC CGAA IAAAGAUU	8745
2904	UUUCUGUC C CCAAUCCC	1346	GGGAUUGG CUGAUGAG GCCGUUAGGC CGAA IACAGAAA	8746
2905	UUCUGUCC C CAAUCCCC	1347	GGGGAUUG CUGAUGAG GCCGUUAGGC CGAA IGACAGAA	8747
2906	UCUGUCCC C AAUCCCCU	1348	AGGGGAUU CUGAUGAG GCCGUUAGGC CGAA IGGACAGA	8748
2907	CUGUCCCC A AUCCCCUG	1349	CAGGGGAU CUGAUGAG GCCGUUAGGC CGAA IGGGACAG	8749
2911	CCCCAAUC C CCUGGGAU	1350	AUCCCAGG CUGAUGAG GCCGUUAGGC CGAA IAUUGGGG	8750
2912	CCCAAUCC C CUGGGAUU	1351	AAUCCCAG CUGAUGAG GCCGUUAGGC CGAA IGAUUGGG	8751
2913	CCAAUCCC C UGGGAUUC	1352	GAAUCCCA CUGAUGAG GCCGUUAGGC CGAA IGGAUUGG	8752
2914	CAAUCCCC U GGGAUUCU	1353	AGAAUCCC CUGAUGAG GCCGUUAGGC CGAA IGGGAUUG	8753
2922	UGGGAUUC U UCCCCGAU	1354	AUCGGGGA CUGAUGAG GCCGUUAGGC CGAA IAAUCCCA	8754
2925	GAUUCUUC C CCGAUCAU	1355	AUGAUCGG CUGAUGAG GCCGUUAGGC CGAA IAAGAAUC	8755
2926	AUUCUUCC C CGAUCAUC	1356	GAUGAUCG CUGAUGAG GCCGUUAGGC CGAA IGAAGAAU	8756
2927	UUCUUCCC C GAUCAUCA	1357	UGAUGAUC CUGAUGAG GCCGUUAGGC CGAA IGGAAGAA	8757
2932	CCCCGAUC A UCAGUUGG	1358	CCAACUGA CUGAUGAG GCCGUUAGGC CGAA IAUCGGGG	8758
2935	CGAUCAUC A GUUGGACC	1359	GGUCCAAC CUGAUGAG GCCGUUAGGC CGAA IAUGAUCG	8759

2943	AGUUGGAC C CUGCAUUC	1360	GAAUGCAG CUGAUGAG GCCGUUAGGC CGAA IUCCAACU	8760
2944	GUUGGACC C UGCAUUCA	1361	UGAAUGCA CUGAUGAG GCCGUUAGGC CGAA IGUCCAAC	8761
2945	UUGGACCC U GCAUUCAA	1362	UUGAAUGC CUGAUGAG GCCGUUAGGC CGAA IGGUCCAA	8762
2948	GACCCUGC A UUCAAAGC	1363	GCUUUGAA CUGAUGAG GCCGUUAGGC CGAA ICAGGGUC	8763
2952	CUGCAUUC A AAGCCAAC	1364	GUUGGCUU CUGAUGAG GCCGUUAGGC CGAA IAAUGCAG	8764
2957	UUCAAAGC C AACUCAGU	1365	ACUGAGUU CUGAUGAG GCCGUUAGGC CGAA ICUUUGAA	8765
2958	UCAAAGCC A ACUCAGUA	1366	UACUGAGU CUGAUGAG GCCGUUAGGC CGAA IGCUUUGA	8766
2961	AAGCCAAC U CAGUAAAU	1367	AUUUACUG CUGAUGAG GCCGUUAGGC CGAA IUUGGCUU	8767
2963	GCCAACUC A GUAAAUCC	1368	GGAUUUAC CUGAUGAG GCCGUUAGGC CGAA IAGUUGGC	8768
2971	AGUAAAUC C AGAUUGGG	1369	CCCAAUCU CUGAUGAG GCCGUUAGGC CGAA IAUUUACU	8769
2972	GUAAAUCC A GAUUGGGA	1370	UCCCAAUC CUGAUGAG GCCGUUAGGC CGAA IGAUUUAC	8770
2982	AUUGGGAC C UCAACCCG	1371	CGGGUUGA CUGAUGAG GCCGUUAGGC CGAA IUCCCAAU	8771
2983	UUGGGACC U CAACCCGC	1372	GCGGGUUG CUGAUGAG GCCGUUAGGC CGAA IGUCCCAA	8772
2985	GGGACCUC A ACCCGCAC	1373	GUGCGGGU CUGAUGAG GCCGUUAGGC CGAA IAGGUCCC	8773
2988	ACCUCAAC C CGCACAAG	1374	CUUGUGCG CUGAUGAG GCCGUUAGGC CGAA IUUGAGGU	8774
2989	CCUCAACC C GCACAAGG	1375	CCUUGUGC CUGAUGAG GCCGUUAGGC CGAA IGUUGAGG	8775
2992	CAACCCGC A CAAGGACA	1376	UGUCCUUG CUGAUGAG GCCGUUAGGC CGAA ICGGGUUG	8776
2994	ACCCGCAC A AGGACAAC	1377	GUUGUCCU CUGAUGAG GCCGUUAGGC CGAA IUGCGGGU	8777
3000	ACAAGGAC A ACUGGCCG	1378	CGGCCAGU CUGAUGAG GCCGUUAGGC CGAA IUCCUUGU	8778
3003	AGGACAAC U GGCCGGAC	1379	GUCCGGCC CUGAUGAG GCCGUUAGGC CGAA IUUGUCCU	8779
3007	CAACUGGC C GGACGCCA	1380	UGGCGUCC CUGAUGAG GCCGUUAGGC CGAA ICCAGUUG	8780
3014	CCGGACGC C AACAAGGU	1381	ACCUUGUU CUGAUGAG GCCGUUAGGC CGAA ICGUCCGG	8781
3015	CGGACGCC A ACAAGGUG	1382	CACCUUGU CUGAUGAG GCCGUUAGGC CGAA IGCGUCCG	8782
3018	ACGCCAAC A AGGUGGGA	1383	UCCCACCU CUGAUGAG GCCGUUAGGC CGAA IUUGGCGU	8783
3035	GUGGGAGC A UUCGGGCC	1384	GGCCCGAA CUGAUGAG GCCGUUAGGC CGAA ICUCCCAC	8784
3043	AUUCGGGC C AGGGUUCA	1385	UGAACCCU CUGAUGAG GCCGUUAGGC CGAA ICCCGAAU	8785
3044	UUCGGGCC A GGGUUCAC	1386	GUGAACCC CUGAUGAG GCCGUUAGGC CGAA IGCCCGAA	8786
3051	CAGGGUUC A CCCCUCCC	1387	GGGAGGGG CUGAUGAG GCCGUUAGGC CGAA IAACCCUG	8787
3053	GGGUUCAC C CCUCCCCA	1388	UGGGGAGG CUGAUGAG GCCGUUAGGC CGAA IUGAACCC	8788
3054	GGUUCACC C CUCCCCAU	1389	AUGGGGAG CUGAUGAG GCCGUUAGGC CGAA IGUGAACC	8789
3055	GUUCACCC C UCCCCAUG	1390	CAUGGGGA CUGAUGAG GCCGUUAGGC CGAA IGGUGAAC	8790
3056	UUCACCCC U CCCCAUGG	1391	CCAUGGGG CUGAUGAG GCCGUUAGGC CGAA IGGGUGAA	8791
3058	CACCCCUC C CCAUGGGG	1392	CCCCAUGG CUGAUGAG GCCGUUAGGC CGAA IAGGGGUG	8792
3059	ACCCCUCC C CAUGGGGG	1393	CCCCCAUG CUGAUGAG GCCGUUAGGC CGAA IGAGGGGU	8793
3060	CCCCUCCC C AUGGGGGA	1394	UCCCCCAU CUGAUGAG GCCGUUAGGC CGAA IGGAGGGG	8794
3061	CCCUCCCC A UGGGGGAC	1395	GUCCCCCA CUGAUGAG GCCGUUAGGC CGAA IGGGAGGG	8795
3070	UGGGGGAC U GUUGGGGU	1396	ACCCCAAC CUGAUGAG GCCGUUAGGC CGAA IUCCCCCA	8796
3084	GGUGGAGC C CUCACGCU	1397	AGCGUGAG CUGAUGAG GCCGUUAGGC CGAA ICUCCACC	8797
3085	GUGGAGCC C UCACGCUC	1398	GAGCGUGA CUGAUGAG GCCGUUAGGC CGAA IGCUCCAC	8798
3086	UGGAGCCC U CACGCUCA	1399	UGAGCGUG CUGAUGAG GCCGUUAGGC CGAA IGGCUCCA	8799
3088	GAGCCCUC A CGCUCAGG	1400	CCUGAGCG CUGAUGAG GCCGUUAGGC CGAA IAGGGCUC	8800
3092	CCUCACGC U CAGGGCCU	1401	AGGCCCUG CUGAUGAG GCCGUUAGGC CGAA ICGUGAGG	8801
3094	UCACGCUC A GGGCCUAC	1402	GUAGGCCC CUGAUGAG GCCGUUAGGC CGAA IAGCGUGA	8802
3099	CUCAGGGC C UACUCACA	1403	UGUGAGUA CUGAUGAG GCCGUUAGGC CGAA ICCCUGAG	8803
3100	UCAGGGCC U ACUCACAA	1404	UUGUGAGU CUGAUGAG GCCGUUAGGC CGAA IGCCCUGA	8804
3103	GGGCCUAC U CACAACUG	1405	CAGUUGUG CUGAUGAG GCCGUUAGGC CGAA IUAGGCCC	8805
3105	GCCUACUC A CAACUGUG	1406	CACAGUUG CUGAUGAG GCCGUUAGGC CGAA IAGUAGGC	8806
3107	CUACUCAC A ACUGUGCC	1407	GGCACAGU CUGAUGAG GCCGUUAGGC CGAA IUGAGUAG	8807
3110	CUCACAAC U GUGCCAGC	1408	GCUGGCAC CUGAUGAG GCCGUUAGGC CGAA IUUGUGAG	8808
3115	AACUGUGC C AGCAGCUC	1409	GAGCUGCU CUGAUGAG GCCGUUAGGC CGAA ICACAGUU	8809
3116	ACUGUGCC A GCAGCUCC	1410	GGAGCUGC CUGAUGAG GCCGUUAGGC CGAA IGCACAGU	8810

3119	GUGCCAGC A GCUCCUCC	1411	GGAGGAGC CUGAUGAG GCCGUUAGGC CGAA ICUGGCAC	8811
3122	CCAGCAGC U CCUCCUCC	1412	GGAGGAGG CUGAUGAG GCCGUUAGGC CGAA ICUGCUGG	8812
3124	AGCAGCUC C UCCUCCUG	1413	CAGGAGGA CUGAUGAG GCCGUUAGGC CGAA IAGCUGCU	8813
3125	GCAGCUCC U CCUCCUGC	1414	GCAGGAGG CUGAUGAG GCCGUUAGGC CGAA IGAGCUGC	8814
3127	AGCUCCUC C UCCUGCCU	1415	AGGCAGGA CUGAUGAG GCCGUUAGGC CGAA IAGGAGCU	8815
3128	GCUCCUCC U CCUGCCUC	1416	GAGGCAGG CUGAUGAG GCCGUUAGGC CGAA IGAGGAGC	8816
3130	UCCUCCUC C UGCCUCCA	1417	UGGAGGCA CUGAUGAG GCCGUUAGGC CGAA IAGGAGGA	8817
3131	CCUCCUCC U GCCUCCAC	1418	GUGGAGGC CUGAUGAG GCCGUUAGGC CGAA IGAGGAGG	8818
3134	CCUCCUGC C UCCACCAA	1419	UUGGUGGA CUGAUGAG GCCGUUAGGC CGAA ICAGGAGG	8819
3135	CUCCUGCC U CCACCAAU	1420	AUUGGUGG CUGAUGAG GCCGUUAGGC CGAA IGCAGGAG	8820
3137	CCUGCCUC C ACCAAUCG	1421	CGAUUGGU CUGAUGAG GCCGUUAGGC CGAA IAGGCAGG	8821
3138	CUGCCUCC A CCAAUCGG	1422	CCGAUUGG CUGAUGAG GCCGUUAGGC CGAA IGAGGCAG	8822
3140	GCCUCCAC C AAUCGGCA	1423	UGCCGAUU CUGAUGAG GCCGUUAGGC CGAA IUGGAGGC	8823
3141	CCUCCACC A AUCGGCAG	1424	CUGCCGAU CUGAUGAG GCCGUUAGGC CGAA IGUGGAGG	8824
3148	CAAUCGGC A GUCAGGAA	1425	UUCCUGAC CUGAUGAG GCCGUUAGGC CGAA ICCGAUUG	8825
3152	CGGCAGUC A GGAAGGCA	1426	UGCCUUCC CUGAUGAG GCCGUUAGGC CGAA IACUGCCG	8826
3160	AGGAAGGC A GCCUACUC	1427	GAGUAGGC CUGAUGAG GCCGUUAGGC CGAA ICCUUCCU	8827
3163	AAGGCAGC C UACUCCCU	1428	AGGGAGUA CUGAUGAG GCCGUUAGGC CGAA ICUGCCUU	8828
3164	AGGCAGCC U ACUCCCUU	1429	AAGGGAGU CUGAUGAG GCCGUUAGGC CGAA IGCUGCCU	8829
3167	CAGCCUAC U CCCUUAUC	1430	GAUAAGGG CUGAUGAG GCCGUUAGGC CGAA IUAGGCUG	8830
3169	GCCUACUC C CUUAUCUC	1431	GAGAUAAG CUGAUGAG GCCGUUAGGC CGAA IAGUAGGC	8831
3170	CCUACUCC C UUAUCUCC	1432	GGAGAUAA CUGAUGAG GCCGUUAGGC CGAA IGAGUAGG	8832
3171	CUACUCCC U UAUCUCCA	1433	UGGAGAUA CUGAUGAG GCCGUUAGGC CGAA IGGAGUAG	8833
3176	CCCUUAUC U CCACCUCU	1434	AGAGGUGG CUGAUGAG GCCGUUAGGC CGAA IAUAAGGG	8834
3178	CUUAUCUC C ACCUCUAA	1435	UUAGAGGU CUGAUGAG GCCGUUAGGC CGAA IAGAUAAG	8835
3179	UUAUCUCC A CCUCUAAG	1436	CUUAGAGG CUGAUGAG GCCGUUAGGC CGAA IGAGAUAA	8836
3181	AUCUCCAC C UCUAAGGG	1437	CCCUUAGA CUGAUGAG GCCGUUAGGC CGAA IUGGAGAU	8837
3182	UCUCCACC U CUAAGGGA	1438	UCCCUUAG CUGAUGAG GCCGUUAGGC CGAA IGUGGAGA	8838
3184	UCCACCUC U AAGGGACA	1439	UGUCCCUU CUGAUGAG GCCGUUAGGC CGAA IAGGUGGA	8839
3192	UAAGGGAC A CUCAUCCU	1440	AGGAUGAG CUGAUGAG GCCGUUAGGC CGAA IUCCCUUA	8840
3194	AGGGACAC U CAUCCUCA	1441	UGAGGAUG CUGAUGAG GCCGUUAGGC CGAA IUGUCCCU	8841
3196	GGACACUC A UCCUCAGG	1442	CCUGAGGA CUGAUGAG GCCGUUAGGC CGAA IAGUGUCC	8842
3199	CACUCAUC C UCAGGCCA	1443	UGGCCUGA CUGAUGAG GCCGUUAGGC CGAA IAUGAGUG	8843
3200	ACUCAUCC U CAGGCCAU	1444	AUGGCCUG CUGAUGAG GCCGUUAGGC CGAA IGAUGAGU	8844
3202	UCAUCCUC A GGCCAUGC	1445	GCAUGGCC CUGAUGAG GCCGUUAGGC CGAA IAGGAUGA	8845
3206	CCUCAGGC C AUGCAGUG	1446	CACUGCAU CUGAUGAG GCCGUUAGGC CGAA ICCUGAGG	8846
3207	CUCAGGCC A UGCAGUGG	1447	CCACUGCA CUGAUGAG GCCGUUAGGC CGAA IGCCUGAG	8847

Input Sequence = AF100308. Cut Site = CH/.
Stem Length = 8 . Core Sequence = CUGAUGAG X CGAA (X = GCCGUUAGGC or other stem II)
AF100308 (Hepatitis B virus strain 2-18, 3215 bp)

Underlined region can be any X sequence or linker, as described herein.
"I" stands for Inosime

TABLE VII: HUMAN HBV G-CLEAVER AND SUBSTRATE SEQUENCE

Pos	Substrate	Seq ID	G-cleaver	Seq ID
61	ACUUUCCU G CUGGUGGC	1448	GCCACCAG UGAUG GCAUGCACUAUGC GCG AGGAAAGU	8848
87	GGAACAGU G AGCCCUGC	1449	GCAGGGCU UGAUG GCAUGCACUAUGC GCG ACUGUUCC	8849
94	UGAGCCCU G CUCAGAAU	1450	AUUCUGAG UGAUG GCAUGCACUAUGC GCG AGGGCUCA	8850
112	CUGUCUCU G CCAUAUCG	1451	CGAUAUGG UGAUG GCAUGCACUAUGC GCG AGAGACAG	8851
132	AUCUUAUC G AAGACUGG	1452	CCAGUCUU UGAUG GCAUGCACUAUGC GCG GAUAAGAU	8852
153	CCUGUACC G AACAUGGA	1453	UCCAUGUU UGAUG GCAUGCACUAUGC GCG GGUACAGG	8853
169	AGAACAUC G CAUCAGGA	1454	UCCUGAUG UGAUG GCAUGCACUAUGC GCG GAUGUUCU	8854
192	GGACCCCU G CUCGUGUU	1455	AACACGAG UGAUG GCAUGCACUAUGC GCG AGGGGUCC	8855
222	UUCUUGUU G ACAAAAAU	1456	AUUUUUGU UGAUG GCAUGCACUAUGC GCG AACAAGAA	8856
315	CAAAAUUC G CAGUCCCA	1457	UGGGACUG UGAUG GCAUGCACUAUGC GCG GAAUUUUG	8857
374	UGGUUAUC G CUGGAUGU	1458	ACAUCCAG UGAUG GCAUGCACUAUGC GCG GAUAACCA	8858
387	AUGUGUCU G CGGCGUUU	1459	AAACGCCG UGAUG GCAUGCACUAUGC GCG AGACACAU	8859
410	CUUCCUCU G CAUCCUGC	1460	GCAGGAUG UGAUG GCAUGCACUAUGC GCG AGAGGAAG	8860
417	UGCAUCCU G CUGCUAUG	1461	CAUAGCAG UGAUG GCAUGCACUAUGC GCG AGGAUGCA	8861
420	AUCCUGCU G CUAUGCCU	1462	AGGCAUAG UGAUG GCAUGCACUAUGC GCG AGCAGGAU	8862
425	GCUGCUAU G CCUCAUCU	1463	AGAUGAGG UGAUG GCAUGCACUAUGC GCG AUAGCAGC	8863
468	GGUAUGUU G CCCGUUUG	1464	CAAACGGG UGAUG GCAUGCACUAUGC GCG AACAUACC	8864
518	CGGACCAU G CAAAACCU	1465	AGGUUUUG UGAUG GCAUGCACUAUGC GCG AUGGUCCG	8865
527	CAAAACCU G CACAACUC	1466	GAGUUGUG UGAUG GCAUGCACUAUGC GCG AGGUUUUG	8866
538	CAACUCCU G CUCAAGGA	1467	UCCUUGAG UGAUG GCAUGCACUAUGC GCG AGGAGUUG	8867
569	CUCAUGUU G CUGUACAA	1468	UUGUACAG UGAUG GCAUGCACUAUGC GCG AACAUGAG	8868
596	CGGAAACU G CACCUGUA	1469	UACAGGUG UGAUG GCAUGCACUAUGC GCG AGUUUCCG	8869
631	GGGCUUUC G CAAAAUAC	1470	GUAUUUUG UGAUG GCAUGCACUAUGC GCG GAAAGCCC	8870
687	UUACUAGU G CCAUUUGU	1471	ACAAAUGG UGAUG GCAUGCACUAUGC GCG ACUAGUAA	8871
747	AUAUGGAU G AUGUGGUU	1472	AACCACAU UGAUG GCAUGCACUAUGC GCG AUCCAUAU	8872
783	AACAUCUU G AGUCCCUU	1473	AAGGGACU UGAUG GCAUGCACUAUGC GCG AAGAUGUU	8873
795	CCCUUUAU G CCGCUGUU	1474	AACAGCGG UGAUG GCAUGCACUAUGC GCG AUAAAGGG	8874
798	UUUAUGCC G CUGUUACC	1475	GGUAACAG UGAUG GCAUGCACUAUGC GCG GGCAUAAA	8875
911	GGCACAUU G CCACAGGA	1476	UCCUGUGG UGAUG GCAUGCACUAUGC GCG AAUGUGCC	8876
978	GGCCUAUU G AUUGGAAA	1477	UUUCCAAU UGAUG GCAUGCACUAUGC GCG AAUAGGCC	8877
997	AUGUCAAC G AAUUGUGG	1478	CCACAAUU UGAUG GCAUGCACUAUGC GCG GUUGACAU	8878
1020	UGGGGUUU G CCGCCCCU	1479	AGGGGCGG UGAUG GCAUGCACUAUGC GCG AAACCCCA	8879
1023	GGUUUGCC G CCCCUUUC	1480	GAAAGGGG UGAUG GCAUGCACUAUGC GCG GGCAAACC	8880
1034	CCUUUCAC G CAAUGUGG	1481	CCACAUUG UGAUG GCAUGCACUAUGC GCG GUGAAAGG	8881
1050	GAUAUUCU G CUUUAAUG	1482	CAUUAAAG UGAUG GCAUGCACUAUGC GCG AGAAUAUC	8882
1058	GCUUUAAU G CCUUUAUA	1483	UAUAAAGG UGAUG GCAUGCACUAUGC GCG AUUAAAGC	8883
1068	CUUUAUAU G CAUGCAUA	1484	UAUGCAUG UGAUG GCAUGCACUAUGC GCG AUAUAAAG	8884
1072	AUAUGCAU G CAUACAAG	1485	CUUGUAUG UGAUG GCAUGCACUAUGC GCG AUGCAUAU	8885
1103	ACUUUCUC G CCAACUUA	1486	UAAGUUGG UGAUG GCAUGCACUAUGC GCG GAGAAAGU	8886
1139	CAGUAUGU G AACCUUUA	1487	UAAAGGUU UGAUG GCAUGCACUAUGC GCG ACAUACUG	8887
1155	ACCCCGUU G CUCGGCAA	1488	UUGCCGAG UGAUG GCAUGCACUAUGC GCG AACGGGGU	8888
1177	UGGUCUAU G CCAAGUGU	1489	ACACUUGG UGAUG GCAUGCACUAUGC GCG AUAGACCA	8889
1188	AAGUGUUU G CUGACGCA	1490	UGCGUCAG UGAUG GCAUGCACUAUGC GCG AAACACUU	8890
1191	UGUUUGCU G ACGCAACC	1491	GGUUGCGU UGAUG GCAUGCACUAUGC GCG AGCAAACA	8891
1194	UUGCUGAC G CAACCCCC	1492	GGGGGUUG UGAUG GCAUGCACUAUGC GCG GUCAGCAA	8892
1234	CCAUCAGC G CAUGCGUG	1493	CACGCAUG UGAUG GCAUGCACUAUGC GCG GCUGAUGG	8893
1238	CAGCGCAU G CGUGGAAC	1494	GUUCCACG UGAUG GCAUGCACUAUGC GCG AUGCGCUG	8894

1262	UCUCCUCU G CCGAUCCA	1495	UGGAUCGG UGAUG GCAUGCACUAUGC GCG AGAGGAGA	8895
1265	CCUCUGCC G AUCCAUAC	1496	GUAUGGAU UGAUG GCAUGCACUAUGC GCG GGCAGAGG	8896
1275	UCCAUACC G CGGAACUC	1497	GAGUUCCG UGAUG GCAUGCACUAUGC GCG GGUAUGGA	8897
1290	UCCUAGCC G CUUGUUUU	1498	AAAACAAG UGAUG GCAUGCACUAUGC GCG GGCUAGGA	8898
1299	CUUGUUUU G CUCGCAGC	1499	GCUGCGAG UGAUG GCAUGCACUAUGC GCG AAAACAAG	8899
1303	UUUUGCUC G CAGCAGGU	1500	ACCUGCUG UGAUG GCAUGCACUAUGC GCG GAGCAAAA	8900
1335	UCGGGACU G ACAAUUCU	1501	AGAAUUGU UGAUG GCAUGCACUAUGC GCG AGUCCCGA	8901
1349	UCUGUCGU G CUCUCCCG	1502	CGGGAGAG UGAUG GCAUGCACUAUGC GCG ACGACAGA	8902
1357	GCUCUCCC G CAAAUAUA	1503	UAUAUUUG UGAUG GCAUGCACUAUGC GCG GGGAGAGC	8903
1382	CCAUGGCU G CUAGGCUG	1504	CAGCCUAG UGAUG GCAUGCACUAUGC GCG AGCCAUGG	8904
1392	UAGGCUGU G CUGCCAAC	1505	GUUGGCAG UGAUG GCAUGCACUAUGC GCG ACAGCCUA	8905
1395	GCUGUGCU G CCAACUGG	1506	CCAGUUGG UGAUG GCAUGCACUAUGC GCG AGCACAGC	8906
1411	GAUCCUAC G CGGGACGU	1507	ACGUCCCG UGAUG GCAUGCACUAUGC GCG GUAGGAUC	8907
1442	CCGUCGGC G CUGAAUCC	1508	GGAUUCAG UGAUG GCAUGCACUAUGC GCG GCCGACGG	8908
1445	UCGGCGCU G AAUCCCGC	1509	GCGGGAUU UGAUG GCAUGCACUAUGC GCG AGCGCCGA	8909
1452	UGAAUCCC G CGGACGAC	1510	GUCGUCCG UGAUG GCAUGCACUAUGC GCG GGGAUUCA	8910
1458	CCGCGGAC G ACCCCUCC	1511	GGAGGGGU UGAUG GCAUGCACUAUGC GCG GUCCGCGG	8911
1474	CCGGGGCC G CUUGGGGC	1512	GCCCCAAG UGAUG GCAUGCACUAUGC GCG GGCCCCGG	8912
1489	GCUCUACC G CCCGCUUC	1513	GAAGCGGG UGAUG GCAUGCACUAUGC GCG GGUAGAGC	8913
1493	UACCGCCC G CUUCUCCG	1514	CGGAGAAG UGAUG GCAUGCACUAUGC GCG GGGCGGUA	8914
1501	GCUUCUCC G CCUAUUGU	1515	ACAAUAGG UGAUG GCAUGCACUAUGC GCG GGAGAAGC	8915
1513	AUUGUACC G ACCGUCCA	1516	UGGACGGU UGAUG GCAUGCACUAUGC GCG GGUACAAU	8916
1528	CACGGGGC G CACCUCUC	1517	GAGAGGUG UGAUG GCAUGCACUAUGC GCG GCCCCGUG	8917
1542	CUCUUUAC G CGGACUCC	1518	GGAGUCCG UGAUG GCAUGCACUAUGC GCG GUAAAGAG	8918
1559	CCGUCUGU G CCUUCUCA	1519	UGAGAAGG UGAUG GCAUGCACUAUGC GCG ACAGACGG	8919
1571	UCUCAUCU G CCGGACCG	1520	CGGUCCGG UGAUG GCAUGCACUAUGC GCG AGAUGAGA	8920
1583	GACCGUGU G CACUUCGC	1521	GCGAAGUG UGAUG GCAUGCACUAUGC GCG ACACGGUC	8921
1590	UGCACUUC G CUUCACCU	1522	AGGUGAAG UGAUG GCAUGCACUAUGC GCG GAAGUGCA	8922
1601	UCACCUCU G CACGUCGC	1523	GCGACGUG UGAUG GCAUGCACUAUGC GCG AGAGGUGA	8923
1608	UGCACGUC G CAUGGAGA	1524	UCUCCAUG UGAUG GCAUGCACUAUGC GCG GACGUGCA	8924
1624	ACCACCGU G AACGCCCA	1525	UGGGCGUU UGAUG GCAUGCACUAUGC GCG ACGGUGGU	8925
1628	CCGUGAAC G CCCACAGG	1526	CCUGUGGG UGAUG GCAUGCACUAUGC GCG GUUCACGG	8926
1642	AGGAACCU G CCCAAGGU	1527	ACCUUGGG UGAUG GCAUGCACUAUGC GCG AGGUUCCU	8927
1654	AAGGUCUU G CAUAAGAG	1528	CUCUUAUG UGAUG GCAUGCACUAUGC GCG AAGACCUU	8928
1690	AUGUCAAC G ACCGACCU	1529	AGGUCGGU UGAUG GCAUGCACUAUGC GCG GUUGACAU	8929
1694	CAACGACC G ACCUUGAG	1530	CUCAAGGU UGAUG GCAUGCACUAUGC GCG GGUCGUUG	8930
1700	CCGACCUU G AGGCAUAC	1531	GUAUGCCU UGAUG GCAUGCACUAUGC GCG AAGGUCGG	8931
1730	UGUUUAAU G AGUGGGAG	1532	CUCCCACU UGAUG GCAUGCACUAUGC GCG AUUAAACA	8932
1818	AGCACCAU G CAACUUUU	1533	AAAAGUUG UGAUG GCAUGCACUAUGC GCG AUGGUGCU	8933
1835	UCACCUCU G CCUAAUCA	1534	UGAUUAGG UGAUG GCAUGCACUAUGC GCG AGAGGUGA	8934
1883	CAAGCUGU G CCUUGGGU	1535	ACCCAAGG UGAUG GCAUGCACUAUGC GCG ACAGCUUG	8935
1912	UGGACAUU G ACCCGUAU	1536	AUACGGGU UGAUG GCAUGCACUAUGC GCG AAUGUCCA	8936
1959	UCUUUUUU G CCUUCUGA	1537	UCAGAAGG UGAUG GCAUGCACUAUGC GCG AAAAAAGA	8937
1966	UGCCUUCU G ACUUCUUU	1538	AAAGAAGU UGAUG GCAUGCACUAUGC GCG AGAAGGCA	8938
1985	UUCUAUUC G AGAUCUCC	1539	GGAGAUCU UGAUG GCAUGCACUAUGC GCG GAAUAGAA	8939
1996	AUCUCCUC G ACACCGCC	1540	GGCGGUGU UGAUG GCAUGCACUAUGC GCG GAGGAGAU	8940
2002	UCGACACC G CCUCUGCU	1541	AGCAGAGG UGAUG GCAUGCACUAUGC GCG GGUGUCGA	8941
2008	CCGCCUCU G CUCUGUAU	1542	AUACAGAG UGAUG GCAUGCACUAUGC GCG AGAGGCGG	8942
2092	GUUGGGGU G AGUUGAUG	1543	CAUCAACU UGAUG GCAUGCACUAUGC GCG ACCCCAAC	8943
2097	GGUGAGUU G AUGAAUCU	1544	AGAUUCAU UGAUG GCAUGCACUAUGC GCG AACUCACC	8944
2100	GAGUUGAU G AAUCUAGC	1545	GCUAGAUU UGAUG GCAUGCACUAUGC GCG AUCAACUC	8945

2237	UUUUGGGC G AGAAACUG	1546	CAGUUUCU UGAUG GCAUGCACUAUGC GCG GCCCAAAA	8946
2251	CUGUUCUU G AAUAUUUG	1547	CAAAUAUU UGAUG GCAUGCACUAUGC GCG AAGAACAG	8947
2282	GUGGAUUC G CACUCCUC	1548	GAGGAGUG UGAUG GCAUGCACUAUGC GCG GAAUCCAC	8948
2293	CUCCUCCU G CAUAUAGA	1549	UCUAUAUG UGAUG GCAUGCACUAUGC GCG AGGAGGAG	8949
2311	CACCAAAU G CCCCUAUC	1550	GAUAGGGG UGAUG GCAUGCACUAUGC GCG AUUUGGUG	8950
2354	UGUUAGAC G AAGAGGCA	1551	UGCCUCUU UGAUG GCAUGCACUAUGC GCG GUCUAACA	8951
2388	ACUCCCUC G CCUCGCAG	1552	CUGCGAGG UGAUG GCAUGCACUAUGC GCG GAGGGAGU	8952
2393	CUCGCCUC G CAGACGAA	1553	UUCGUCUG UGAUG GCAUGCACUAUGC GCG GAGGCGAG	8953
2399	UCGCAGAC G AAGGUCUC	1554	GAGACCUU UGAUG GCAUGCACUAUGC GCG GUCUGCGA	8954
2412	UCUCAAUC G CCGCGUCG	1555	CGACGCGG UGAUG GCAUGCACUAUGC GCG GAUUGAGA	8955
2415	CAAUCGCC G CGUCGCAG	1556	CUGCGACG UGAUG GCAUGCACUAUGC GCG GGCGAUUG	8956
2420	GCCGCGUC G CAGAAGAU	1557	AUCUUCUG UGAUG GCAUGCACUAUGC GCG GACGCGGC	8957
2514	GGUACCUU G CUUUAAUC	1558	GAUUAAAG UGAUG GCAUGCACUAUGC GCG AAGGUACC	8958
2549	CUUUUCCU G ACAUUCAU	1559	AUGAAUGU UGAUG GCAUGCACUAUGC GCG AGGAAAAG	8959
2560	AUUCAUUU G CAGGAGGA	1560	UCCUCCUG UGAUG GCAUGCACUAUGC GCG AAAUGAAU	8960
2576	ACAUUGUU G AUAGAUGU	1561	ACAUCUAU UGAUG GCAUGCACUAUGC GCG AACAAUGU	8961
2615	CAGUAAAU G AAAACAGG	1562	CCUGUUUU UGAUG GCAUGCACUAUGC GCG AUUUACUG	8962
2641	UUAACUAU G CCUGCUAG	1563	CUAGCAGG UGAUG GCAUGCACUAUGC GCG AUAGUUAA	8963
2645	CUAUGCCU G CUAGGUUU	1564	AAACCUAG UGAUG GCAUGCACUAUGC GCG AGGCAUAG	8964
2677	AAAUAUUU G CCCUUAGA	1565	UCUAAGGG UGAUG GCAUGCACUAUGC GCG AAAUAUUU	8965
2740	UUCCAGAC G CGACAUUA	1566	UAAUGUCG UGAUG GCAUGCACUAUGC GCG GUCUGGAA	8966
2742	CCAGACGC G ACAUUAUU	1567	AAUAAUGU UGAUG GCAUGCACUAUGC GCG GCGUCUGG	8967
2804	CACGUAGC G CCUCAUUU	1568	AAAUGAGG UGAUG GCAUGCACUAUGC GCG GCUACGUG	8968
2814	CUCAUUUU G CGGGUCAC	1569	GUGACCCG UGAUG GCAUGCACUAUGC GCG AAAAUGAG	8969
2875	CAAACCUC G AAAAGGCA	1570	UGCCUUUU UGAUG GCAUGCACUAUGC GCG GAGGUUUG	8970
2928	UCUUCCCC G AUCAUCAG	1571	CUGAUGAU UGAUG GCAUGCACUAUGC GCG GGGGAAGA	8971
2946	UGGACCCU G CAUUCAAA	1572	UUUGAAUG UGAUG GCAUGCACUAUGC GCG AGGGUCCA	8972
2990	CUCAACCC G CACAAGGA	1573	UCCUUGUG UGAUG GCAUGCACUAUGC GCG GGGUUGAG	8973
3012	GGCCGGAC G CCAACAAG	1574	CUUGUUGG UGAUG GCAUGCACUAUGC GCG GUCCGGCC	8974
3090	GCCCUCAC G CUCAGGGC	1575	GCCCUGAG UGAUG GCAUGCACUAUGC GCG GUGAGGGC	8975
3113	ACAACUGU G CCAGCAGC	1576	GCUGCUGG UGAUG GCAUGCACUAUGC GCG ACAGUUGU	8976
3132	CUCCUCCU G CCUCCACC	1577	GGUGGAGG UGAUG GCAUGCACUAUGC GCG AGGAGGAG	8977
51	AGGGCCCU G UACUUUCC	1578	GGAAAGUA UGAUG GCAUGCACUAUGC GCG AGGGCCCU	8978
106	AGAAUACU G UCUCUGCC	1579	GGCAGAGA UGAUG GCAUGCACUAUGC GCG AGUAUUCU	8979
148	GGGACCCU G UACCGAAC	1580	GUUCGGUA UGAUG GCAUGCACUAUGC GCG AGGGUCCC	8980
198	CUGCUCGU G UUACAGGC	1581	GCCUGUAA UGAUG GCAUGCACUAUGC GCG ACGAGCAG	8981
219	UUUUUCUU G UUGACAAA	1582	UUUGUCAA UGAUG GCAUGCACUAUGC GCG AAGAAAAA	8982
297	ACACCCGU G UGUCUUGG	1583	CCAAGACA UGAUG GCAUGCACUAUGC GCG ACGGGUGU	8983
299	ACCCGUGU G UCUUGGCC	1584	GGCCAAGA UGAUG GCAUGCACUAUGC GCG ACACGGGU	8984
347	ACCAACCU G UUGUCCUC	1585	GAGGACAA UGAUG GCAUGCACUAUGC GCG AGGUUGGU	8985
350	AACCUGUU G UCCUCCAA	1586	UUGGAGGA UGAUG GCAUGCACUAUGC GCG AACAGGUU	8986
362	UCCAAUUU G UCCUGGUU	1587	AACCAGGA UGAUG GCAUGCACUAUGC GCG AAAUUGGA	8987
381	CGCUGGAU G UGUCUGCG	1588	CGCAGACA UGAUG GCAUGCACUAUGC GCG AUCCAGCG	8988
383	CUGGAUGU G UCUGCGGC	1589	GCCGCAGA UGAUG GCAUGCACUAUGC GCG ACAUCCAG	8989
438	AUCUUCUU G UUGGUUCU	1590	AGAACCAA UGAUG GCAUGCACUAUGC GCG AAGAAGAU	8990
465	CAAGGUAU G UUGCCCGU	1591	ACGGGCAA UGAUG GCAUGCACUAUGC GCG AUACCUUG	8991
476	GCCCGUUU G UCCUCUAA	1592	UUAGAGGA UGAUG GCAUGCACUAUGC GCG AAACGGGC	8992
555	ACCUCUAU G UUUCCCUC	1593	GAGGGAAA UGAUG GCAUGCACUAUGC GCG AUAGAGGU	8993
566	UCCCUCAU G UUGCUGUA	1594	UACAGCAA UGAUG GCAUGCACUAUGC GCG AUGAGGGA	8994
572	AUGUUGCU G UACAAAAC	1595	GUUUUGUA UGAUG GCAUGCACUAUGC GCG AGCAACAU	8995
602	CUGCACCU G UAUUCCCA	1596	UGGGAAUA UGAUG GCAUGCACUAUGC GCG AGGUGCAG	8996

694	UGCCAUUU G UUCAGUGG	1597	CCACUGAA UGAUG GCAUGCACUAUGC GCG AAAUGGCA	8997
724	CCCCCACU G UCUGGCUU	1598	AAGCCAGA UGAUG GCAUGCACUAUGC GCG AGUGGGGG	8998
750	UGGAUGAU G UGGUUUUG	1599	CAAAACCA UGAUG GCAUGCACUAUGC GCG AUCAUCCA	8999
771	CCAAGUCU G UACAACAU	1600	AUGUUGUA UGAUG GCAUGCACUAUGC GCG AGACUUGG	9000
801	AUGCCGCU G UUACCAAU	1601	AUUGGUAA UGAUG GCAUGCACUAUGC GCG AGCGGCAU	9001
818	UUUCUUUU G UCUUUGGG	1602	CCCAAAGA UGAUG GCAUGCACUAUGC GCG AAAAGAAA	9002
888	UGGGAUAU G UAAUUGGG	1603	CCCAAUUA UGAUG GCAUGCACUAUGC GCG AUAUCCCA	9003
927	AACAUAUU G UACAAAAA	1604	UUUUUGUA UGAUG GCAUGCACUAUGC GCG AAUAUGUU	9004
944	AUCAAAAU G UGUUUUAG	1605	CUAAAACA UGAUG GCAUGCACUAUGC GCG AUUUUGAU	9005
946	CAAAAUGU G UUUUAGGA	1606	UCCUAAAA UGAUG GCAUGCACUAUGC GCG ACAUUUUG	9006
963	AACUUCCU G UAAACAGG	1607	CCUGUUUA UGAUG GCAUGCACUAUGC GCG AGGAAGUU	9007
991	GAAAGUAU G UCAACGAA	1608	UUCGUUGA UGAUG GCAUGCACUAUGC GCG AUACUUUC	9008
1002	AACGAAUU G UGGGUCUU	1609	AAGACCCA UGAUG GCAUGCACUAUGC GCG AAUUCGUU	9009
1039	CACGCAAU G UGGAUAUU	1610	AAUAUCCA UGAUG GCAUGCACUAUGC GCG AUUGCGUG	9010
1137	AACAGUAU G UGAACCUU	1611	AAGGUUCA UGAUG GCAUGCACUAUGC GCG AUACUGUU	9011
1184	UGCCAAGU G UUUGCUGA	1612	UCAGCAAA UGAUG GCAUGCACUAUGC GCG ACUUGGCA	9012
1251	GAACCUUU G UGUCUCCU	1613	AGGAGACA UGAUG GCAUGCACUAUGC GCG AAAGGUUC	9013
1253	ACCUUUGU G UCUCCUCU	1614	AGAGGAGA UGAUG GCAUGCACUAUGC GCG ACAAAGGU	9014
1294	AGCCGCUU G UUUUGCUC	1615	GAGCAAAA UGAUG GCAUGCACUAUGC GCG AAGCGGCU	9015
1344	ACAAUUCU G UCGUGCUC	1616	GAGCACGA UGAUG GCAUGCACUAUGC GCG AGAAUUGU	9016
1390	GCUAGGCU G UGCUGCCA	1617	UGGCAGCA UGAUG GCAUGCACUAUGC GCG AGCCUAGC	9017
1425	CGUCCUUU G UUUACGUC	1618	GACGUAAA UGAUG GCAUGCACUAUGC GCG AAAGGACG	9018
1508	CGCCUAUU G UACCGACC	1619	GGUCGGUA UGAUG GCAUGCACUAUGC GCG AAUAGGCG	9019
1557	CCCCGUCU G UGCCUUCU	1620	AGAAGGCA UGAUG GCAUGCACUAUGC GCG AGACGGGG	9020
1581	CGGACCGU G UGCACUUC	1621	GAAGUGCA UGAUG GCAUGCACUAUGC GCG ACGGUCCG	9021
1684	UCAGCAAU G UCAACGAC	1622	GUCGUUGA UGAUG GCAUGCACUAUGC GCG AUUGCUGA	9022
1719	CAAAGACU G UGUGUUUA	1623	UAAACACA UGAUG GCAUGCACUAUGC GCG AGUCUUUG	9023
1721	AAGACUGU G UGUUUAAU	1624	AUUAAACA UGAUG GCAUGCACUAUGC GCG ACAGUCUU	9024
1723	GACUGUGU G UUUAAUGA	1625	UCAUUAAA UGAUG GCAUGCACUAUGC GCG ACACAGUC	9025
1772	AGGUCUUU G UACUAGGA	1626	UCCUAGUA UGAUG GCAUGCACUAUGC GCG AAAGACCU	9026
1785	AGGAGGCU G UAGGCAUA	1627	UAUGCCUA UGAUG GCAUGCACUAUGC GCG AGCCUCCU	9027
1801	AAAUUGGU G UGUUCACC	1628	GGUGAACA UGAUG GCAUGCACUAUGC GCG ACCAAUUU	9028
1803	AUUGGUGU G UUCACCAG	1629	CUGGUGAA UGAUG GCAUGCACUAUGC GCG ACACCAAU	9029
1850	CAUCUCAU G UUCAUGUC	1630	GACAUGAA UGAUG GCAUGCACUAUGC GCG AUGAGAUG	9030
1856	AUGUUCAU G UCCUACUG	1631	CAGUAGGA UGAUG GCAUGCACUAUGC GCG AUGAACAU	9031
1864	GUCCUACU G UUCAAGCC	1632	GGCUUGAA UGAUG GCAUGCACUAUGC GCG AGUAGGAC	9032
1881	UCCAAGCU G UGCCUUGG	1633	CCAAGGCA UGAUG GCAUGCACUAUGC GCG AGCUUGGA	9033
1939	GAGCUUCU G UGGAGUUA	1634	UAACUCCA UGAUG GCAUGCACUAUGC GCG AGAAGCUC	9034
2013	UCUGCUCU G UAUCGGGG	1635	CCCCGAUA UGAUG GCAUGCACUAUGC GCG AGAGCAGA	9035
2045	GGAACAUU G UUCACCUC	1636	GAGGUGAA UGAUG GCAUGCACUAUGC GCG AAUGUUCC	9036
2082	GCUAUUCU G UGUUGGGG	1637	CCCCAACA UGAUG GCAUGCACUAUGC GCG AGAAUAGC	9037
2084	UAUUCUGU G UUGGGGUG	1638	CACCCCAA UGAUG GCAUGCACUAUGC GCG ACAGAAUA	9038
2167	UCAGCUAU G UCAACGUU	1639	AACGUUGA UGAUG GCAUGCACUAUGC GCG AUAGCUGA	9039
2205	CAACUAUU G UGGUUUCA	1640	UGAAACCA UGAUG GCAUGCACUAUGC GCG AAUAGUUG	9040
2222	CAUUUCCU G UCUUACUU	1641	AAGUAAGA UGAUG GCAUGCACUAUGC GCG AGGAAAUG	9041
2245	GAGAAACU G UUCUUGAA	1642	UUCAAGAA UGAUG GCAUGCACUAUGC GCG AGUUUCUC	9042
2262	UAUUUGGU G UCUUUUGG	1643	CCAAAAGA UGAUG GCAUGCACUAUGC GCG ACCAAAUA	9043
2274	UUUGGAGU G UGGAUUCG	1644	CGAAUCCA UGAUG GCAUGCACUAUGC GCG ACUCCAAA	9044
2344	AAACUACU G UUGUUAGA	1645	UCUAACAA UGAUG GCAUGCACUAUGC GCG AGUAGUUU	9045
2347	CUACUGUU G UUAGACGA	1646	UCGUCUAA UGAUG GCAUGCACUAUGC GCG AACAGUAG	9046
2450	AUCUCAAU G UUAGUAUU	1647	AAUACUAA UGAUG GCAUGCACUAUGC GCG AUUGAGAU	9047

2573	AGGACAUU G UUGAUAGA	1648	UCUAUCAA UGAUG GCAUGCACUAUGC GCG AAUGUCCU	9048
2583	UGAUAGAU G UAAGCAAU	1649	AUUGCUUA UGAUG GCAUGCACUAUGC GCG AUCUAUCA	9049
2594	AGCAAUUU G UGGGGCCC	1650	GGGCCCCA UGAUG GCAUGCACUAUGC GCG AAAUUGCU	9050
2663	AUCCCAAU G UUACUAAA	1651	UUUAGUAA UGAUG GCAUGCACUAUGC GCG AUUGGGAU	9051
2717	CAGAGUAU G UAGUUAAU	1652	AUUAACUA UGAUG GCAUGCACUAUGC GCG AUACUCUG	9052
2901	AUCUUUCU G UCCCCAAU	1653	AUUGGGGA UGAUG GCAUGCACUAUGC GCG AGAAAGAU	9053
3071	GGGGGACU G UUGGGGUG	1654	CACCCCAA UGAUG GCAUGCACUAUGC GCG AGUCCCCC	9054
3111	UCACAACU G UGCCAGCA	1655	UGCUGGCA UGAUG GCAUGCACUAUGC GCG AGUUGUGA	9055

Input Sequence = AF100308. Cut Site = YG/M or UG/U.
Stem Length = 8. Core Sequence = UGAUG GCAUGCACUAUGC GCG
AF100308 (Hepatitis B virus strain 2-18, 3215 bp)

TABLE VIII: HUMAN HBV ZINZYME AND SUBSTRATE SEQUENCE

Pos	Substrate	Seq ID	Zinzyme	Seq ID
61	ACUUUCCU G CUGGUGGC	1448	GCCACCAG GCcgaaagGCGaGuCaaGGuCu AGGAAAGU	9056
94	UGAGCCCU G CUCAGAAU	1450	AUUCUGAG GCcgaaagGCGaGuCaaGGuCu AGGGCUCA	9057
112	CUGUCUCU G CCAUAUCG	1451	CGAUAUGG GCcgaaagGCGaGuCaaGGuCu AGAGACAG	9058
169	AGAACAUC G CAUCAGGA	1454	UCCUGAUG GCcgaaagGCGaGuCaaGGuCu GAUGUUCU	9059
192	GGACCCCU G CUCGUGUU	1455	AACACGAG GCcgaaagGCGaGuCaaGGuCu AGGGGUCC	9060
315	CAAAAUUC G CAGUCCCA	1457	UGGGACUG GCcgaaagGCGaGuCaaGGuCu GAAUUUUG	9061
374	UGGUUAUC G CUGGAUGU	1458	ACAUCCAG GCcgaaagGCGaGuCaaGGuCu GAUAACCA	9062
387	AUGUGUCU G CGGCGUUU	1459	AAACGCCG GCcgaaagGCGaGuCaaGGuCu AGACACAU	9063
410	CUUCCUCU G CAUCCUGC	1460	GCAGGAUG GCcgaaagGCGaGuCaaGGuCu AGAGGAAG	9064
417	UGCAUCCU G CUGCUAUG	1461	CAUAGCAG GCcgaaagGCGaGuCaaGGuCu AGGAUGCA	9065
420	AUCCUGCU G CUAUGCCU	1462	AGGCAUAG GCcgaaagGCGaGuCaaGGuCu AGCAGGAU	9066
425	GCUGCUAU G CCUCAUCU	1463	AGAUGAGG GCcgaaagGCGaGuCaaGGuCu AUAGCAGC	9067
468	GGUAUGUU G CCCGUUUG	1464	CAAACGGG GCcgaaagGCGaGuCaaGGuCu AACAUACC	9068
518	CGGACCAU G CAAAACCU	1465	AGGUUUUG GCcgaaagGCGaGuCaaGGuCu AUGGUCCG	9069
527	CAAAACCU G CACAACUC	1466	GAGUUGUG GCcgaaagGCGaGuCaaGGuCu AGGUUUUG	9070
538	CAACUCCU G CUCAAGGA	1467	UCCUUGAG GCcgaaagGCGaGuCaaGGuCu AGGAGUUG	9071
569	CUCAUGUU G CUGUACAA	1468	UUGUACAG GCcgaaagGCGaGuCaaGGuCu AACAUGAG	9072
596	CGGAAACU G CACCUGUA	1469	UACAGGUG GCcgaaagGCGaGuCaaGGuCu AGUUUCCG	9073
631	GGGCUUUC G CAAAAUAC	1470	GUAUUUUG GCcgaaagGCGaGuCaaGGuCu GAAAGCCC	9074
687	UUACUAGU G CCAUUUGU	1471	ACAAAUGG GCcgaaagGCGaGuCaaGGuCu ACUAGUAA	9075
795	CCCUUUAU G CCGCUGUU	1474	AACAGCGG GCcgaaagGCGaGuCaaGGuCu AUAAAGGG	9076
798	UUUAUGCC G CUGUUACC	1475	GGUAACAG GCcgaaagGCGaGuCaaGGuCu GGCAUAAA	9077
911	GGCACAUU G CCACAGGA	1476	UCCUGUGG GCcgaaagGCGaGuCaaGGuCu AAUGUGCC	9078
1020	UGGGGUUU G CCGCCCCU	1479	AGGGGCGG GCcgaaagGCGaGuCaaGGuCu AAACCCCA	9079
1023	GGUUUGCC G CCCCUUUC	1480	GAAAGGGG GCcgaaagGCGaGuCaaGGuCu GGCAAACC	9080
1034	CCUUUCAC G CAAUGUGG	1481	CCACAUUG GCcgaaagGCGaGuCaaGGuCu GUGAAAGG	9081
1050	GAUAUUCU G CUUUAAUG	1482	CAUUAAAG GCcgaaagGCGaGuCaaGGuCu AGAAUAUC	9082
1058	GCUUUAAU G CCUUUAUA	1483	UAUAAAGG GCcgaaagGCGaGuCaaGGuCu AUUAAAGC	9083
1068	CUUUAUAU G CAUGCAUA	1484	UAUGCAUG GCcgaaagGCGaGuCaaGGuCu AUAUAAAG	9084
1072	AUAUGCAU G CAUACAAG	1485	CUUGUAUG GCcgaaagGCGaGuCaaGGuCu AUGCAUAU	9085
1103	ACUUUCUC G CCAACUUA	1486	UAAGUUGG GCcgaaagGCGaGuCaaGGuCu GAGAAAGU	9086
1155	ACCCCGUU G CUCGGCAA	1488	UUGCCGAG GCcgaaagGCGaGuCaaGGuCu AACGGGGU	9087
1177	UGGUCUAU G CCAAGUGU	1489	ACACUUGG GCcgaaagGCGaGuCaaGGuCu AUAGACCA	9088
1188	AAGUGUUU G CUGACGCA	1490	UGCGUCAG GCcgaaagGCGaGuCaaGGuCu AAACACUU	9089
1194	UUGCUGAC G CAACCCCC	1492	GGGGGUUG GCcgaaagGCGaGuCaaGGuCu GUCAGCAA	9090
1234	CCAUCAGC G CAUGCGUG	1493	CACGCAUG GCcgaaagGCGaGuCaaGGuCu GCUGAUGG	9091
1238	CAGCGCAU G CGUGGAAC	1494	GUUCCACG GCcgaaagGCGaGuCaaGGuCu AUGCGCUG	9092
1262	UCUCCUCU G CCGAUCCA	1495	UGGAUCGG GCcgaaagGCGaGuCaaGGuCu AGAGGAGA	9093
1275	UCCAUACC G CGGAACUC	1497	GAGUUCCG GCcgaaagGCGaGuCaaGGuCu GGUAUGGA	9094
1290	UCCUAGCC G CUUGUUUU	1498	AAAACAAG GCcgaaagGCGaGuCaaGGuCu GGCUAGGA	9095
1299	CUUGUUUU G CUCGCAGC	1499	GCUGCGAG GCcgaaagGCGaGuCaaGGuCu AAAACAAG	9096
1303	UUUUGCUC G CAGCAGGU	1500	ACCUGCUG GCcgaaagGCGaGuCaaGGuCu GAGCAAAA	9097
1349	UCUGUCGU G CUCUCCCG	1502	CGGGAGAG GCcgaaagGCGaGuCaaGGuCu ACGACAGA	9098
1357	GCUCUCCC G CAAAUAUA	1503	UAUAUUUG GCcgaaagGCGaGuCaaGGuCu GGGAGAGC	9099
1382	CCAUGGCU G CUAGGCUG	1504	CAGCCUAG GCcgaaagGCGaGuCaaGGuCu AGCCAUGG	9100
1392	UAGGCUGU G CUGCCAAC	1505	GUUGGCAG GCcgaaagGCGaGuCaaGGuCu ACAGCCUA	9101
1395	GCUGUGCU G CCAACUGG	1506	CCAGUUGG GCcgaaagGCGaGuCaaGGuCu AGCACAGC	9102

1411	GAUCCUAC G CGGGACGU	1507	ACGUCCCG GCcgaaagGCGaGuCaaGGuCu GUAGGAUC	9103
1442	CCGUCGGC G CUGAAUCC	1508	GGAUUCAG GCcgaaagGCGaGuCaaGGuCu GCCGACGG	9104
1452	UGAAUCCC G CGGACGAC	1510	GUCGUCCG GCcgaaagGCGaGuCaaGGuCu GGGAUUCA	9105
1474	CCGGGGCC G CUUGGGGC	1512	GCCCCAAG GCcgaaagGCGaGuCaaGGuCu GGCCCCGG	9106
1489	GCUCUACC G CCCGCUUC	1513	GAAGCGGG GCcgaaagGCGaGuCaaGGuCu GGUAGAGC	9107
1493	UACCGCCC G CUUCUCCG	1514	CGGAGAAG GCcgaaagGCGaGuCaaGGuCu GGGCGGUA	9108
1501	GCUUCUCC G CCUAUUGU	1515	ACAAUAGG GCcgaaagGCGaGuCaaGGuCu GGAGAAGC	9109
1528	CACGGGGC G CACCUCUC	1517	GAGAGGUG GCcgaaagGCGaGuCaaGGuCu GCCCCGUG	9110
1542	CUCUUUAC G CGGACUCC	1518	GGAGUCCG GCcgaaagGCGaGuCaaGGuCu GUAAAGAG	9111
1559	CCGUCUGU G CCUUCUCA	1519	UGAGAAGG GCcgaaagGCGaGuCaaGGuCu ACAGACGG	9112
1571	UCUCAUCU G CCGGACCG	1520	CGGUCCGG GCcgaaagGCGaGuCaaGGuCu AGAUGAGA	9113
1583	GACCGUGU G CACUUCGC	1521	GCGAAGUG GCcgaaagGCGaGuCaaGGuCu ACACGGUC	9114
1590	UGCACUUC G CUUCACCU	1522	AGGUGAAG GCcgaaagGCGaGuCaaGGuCu GAAGUGCA	9115
1601	UCACCUCU G CACGUCGC	1523	GCGACGUG GCcgaaagGCGaGuCaaGGuCu AGAGGUGA	9116
1608	UGCACGUC G CAUGGAGA	1524	UCUCCAUG GCcgaaagGCGaGuCaaGGuCu GACGUGCA	9117
1628	CCGUGAAC G CCCACAGG	1526	CCUGUGGG GCcgaaagGCGaGuCaaGGuCu GUUCACGG	9118
1642	AGGAACCU G CCCAAGGU	1527	ACCUUGGG GCcgaaagGCGaGuCaaGGuCu AGGUUCCU	9119
1654	AAGGUCUU G CAUAAGAG	1528	CUCUUAUG GCcgaaagGCGaGuCaaGGuCu AAGACCUU	9120
1818	AGCACCAU G CAACUUUU	1533	AAAAGUUG GCcgaaagGCGaGuCaaGGuCu AUGGUGCU	9121
1835	UCACCUCU G CCUAAUCA	1534	UGAUUAGG GCcgaaagGCGaGuCaaGGuCu AGAGGUGA	9122
1883	CAAGCUGU G CCUUGGGU	1535	ACCCAAGG GCcgaaagGCGaGuCaaGGuCu ACAGCUUG	9123
1959	UCUUUUUU G CCUUCUGA	1537	UCAGAAGG GCcgaaagGCGaGuCaaGGuCu AAAAAAGA	9124
2002	UCGACACC G CCUCUGCU	1541	AGCAGAGG GCcgaaagGCGaGuCaaGGuCu GGUGUCGA	9125
2008	CCGCCUCU G CUCUGUAU	1542	AUACAGAG GCcgaaagGCGaGuCaaGGuCu AGAGGCGG	9126
2282	GUGGAUUC G CACUCCUC	1548	GAGGAGUG GCcgaaagGCGaGuCaaGGuCu GAAUCCAC	9127
2293	CUCCUCCU G CAUAUAGA	1549	UCUAUAUG GCcgaaagGCGaGuCaaGGuCu AGGAGGAG	9128
2311	CACCAAAU G CCCCUAUC	1550	GAUAGGGG GCcgaaagGCGaGuCaaGGuCu AUUUGGUG	9129
2388	ACUCCCUC G CCUCGCAG	1552	CUGCGAGG GCcgaaagGCGaGuCaaGGuCu GAGGGAGU	9130
2393	CUCGCCUC G CAGACGAA	1553	UUCGUCUG GCcgaaagGCGaGuCaaGGuCu GAGGCGAG	9131
2412	UCUCAAUC G CCGCGUCG	1555	CGACGCGG GCcgaaagGCGaGuCaaGGuCu GAUUGAGA	9132
2415	CAAUCGCC G CGUCGCAG	1556	CUGCGACG GCcgaaagGCGaGuCaaGGuCu GGCGAUUG	9133
2420	GCCGCGUC G CAGAAGAU	1557	AUCUUCUG GCcgaaagGCGaGuCaaGGuCu GACGCGGC	9134
2514	GGUACCUU G CUUUAAUC	1558	GAUUAAAG GCcgaaagGCGaGuCaaGGuCu AAGGUACC	9135
2560	AUUCAUUU G CAGGAGGA	1560	UCCUCCUG GCcgaaagGCGaGuCaaGGuCu AAAUGAAU	9136
2641	UUAACUAU G CCUGCUAG	1563	CUAGCAGG GCcgaaagGCGaGuCaaGGuCu AUAGUUAA	9137
2645	CUAUGCCU G CUAGGUUU	1564	AAACCUAG GCcgaaagGCGaGuCaaGGuCu AGGCAUAG	9138
2677	AAAUAUUU G CCCUUAGA	1565	UCUAAGGG GCcgaaagGCGaGuCaaGGuCu AAAUAUUU	9139
2740	UUCCAGAC G CGACAUUA	1566	UAAUGUCG GCcgaaagGCGaGuCaaGGuCu GUCUGGAA	9140
2804	CACGUAGC G CCUCAUUU	1568	AAAUGAGG GCcgaaagGCGaGuCaaGGuCu GCUACGUG	9141
2814	CUCAUUUU G CGGGUCAC	1569	GUGACCCG GCcgaaagGCGaGuCaaGGuCu AAAAUGAG	9142
2946	UGGACCCU G CAUUCAAA	1572	UUUGAAUG GCcgaaagGCGaGuCaaGGuCu AGGGUCCA	9143
2990	CUCAACCC G CACAAGGA	1573	UCCUUGUG GCcgaaagGCGaGuCaaGGuCu GGGUUGAG	9144
3012	GGCCGGAC G CCAACAAG	1574	CUUGUUGG GCcgaaagGCGaGuCaaGGuCu GUCCGGCC	9145
3090	GCCCUCAC G CUCAGGGC	1575	GCCCUGAG GCcgaaagGCGaGuCaaGGuCu GUGAGGGC	9146
3113	ACAACUGU G CCAGCAGC	1576	GCUGCUGG GCcgaaagGCGaGuCaaGGuCu ACAGUUGU	9147
3132	CUCCUCCU G CCUCCACC	1577	GGUGGAGG GCcgaaagGCGaGuCaaGGuCu AGGAGGAG	9148
51	AGGGCCCU G UACUUUCC	1578	GGAAAGUA GCcgaaagGCGaGuCaaGGuCu AGGGCCCU	9149
106	AGAAUACU G UCUCUGCC	1579	GGCAGAGA GCcgaaagGCGaGuCaaGGuCu AGUAUUCU	9150
148	GGGACCCU G UACCGAAC	1580	GUUCGGUA GCcgaaagGCGaGuCaaGGuCu AGGGUCCC	9151
198	CUGCUCGU G UUACAGGC	1581	GCCUGUAA GCcgaaagGCGaGuCaaGGuCu ACGAGCAG	9152
219	UUUUUCUU G UUGACAAA	1582	UUUGUCAA GCcgaaagGCGaGuCaaGGuCu AAGAAAAA	9153

297	ACACCCGU G UGUCUUGG	1583	CCAAGACA GCcgaaagGCGaGuCaaGGuCu ACGGGUGU	9154
299	ACCCGUGU G UCUUGGCC	1584	GGCCAAGA GCcgaaagGCGaGuCaaGGuCu ACACGGGU	9155
347	ACCAACCU G UUGUCCUC	1585	GAGGACAA GCcgaaagGCGaGuCaaGGuCu AGGUUGGU	9156
350	AACCUGUU G UCCUCCAA	1586	UUGGAGGA GCcgaaagGCGaGuCaaGGuCu AACAGGUU	9157
362	UCCAAUUU G UCCUGGUU	1587	AACCAGGA GCcgaaagGCGaGuCaaGGuCu AAAUUGGA	9158
381	CGCUGGAU G UGUCUGCG	1588	CGCAGACA GCcgaaagGCGaGuCaaGGuCu AUCCAGCG	9159
383	CUGGAUGU G UCUGCGGC	1589	GCCGCAGA GCcgaaagGCGaGuCaaGGuCu ACAUCCAG	9160
438	AUCUUCUU G UUGGUUCU	1590	AGAACCAA GCcgaaagGCGaGuCaaGGuCu AAGAAGAU	9161
465	CAAGGUAU G UUGCCCGU	1591	ACGGGCAA GCcgaaagGCGaGuCaaGGuCu AUACCUUG	9162
476	GCCCGUUU G UCCUCUAA	1592	UUAGAGGA GCcgaaagGCGaGuCaaGGuCu AAACGGGC	9163
555	ACCUCUAU G UUUCCCUC	1593	GAGGGAAA GCcgaaagGCGaGuCaaGGuCu AUAGAGGU	9164
566	UCCCUCAU G UUGCUGUA	1594	UACAGCAA GCcgaaagGCGaGuCaaGGuCu AUGAGGGA	9165
572	AUGUUGCU G UACAAAAC	1595	GUUUUGUA GCcgaaagGCGaGuCaaGGuCu AGCAACAU	9166
602	CUGCACCU G UAUUCCCA	1596	UGGGAAUA GCcgaaagGCGaGuCaaGGuCu AGGUGCAG	9167
694	UGCCAUUU G UUCAGUGG	1597	CCACUGAA GCcgaaagGCGaGuCaaGGuCu AAAUGGCA	9168
724	CCCCCACU G UCUGGCUU	1598	AAGCCAGA GCcgaaagGCGaGuCaaGGuCu AGUGGGGG	9169
750	UGGAUGAU G UGGUUUUG	1599	CAAAACCA GCcgaaagGCGaGuCaaGGuCu AUCAUCCA	9170
771	CCAAGUCU G UACAACAU	1600	AUGUUGUA GCcgaaagGCGaGuCaaGGuCu AGACUUGG	9171
801	AUGCCGCU G UUACCAAU	1601	AUUGGUAA GCcgaaagGCGaGuCaaGGuCu AGCGGCAU	9172
818	UUUCUUUU G UCUUUGGG	1602	CCCAAAGA GCcgaaagGCGaGuCaaGGuCu AAAAGAAA	9173
888	UGGGAUAU G UAAUUGGG	1603	CCCAAUUA GCcgaaagGCGaGuCaaGGuCu AUAUCCCA	9174
927	AACAUAUU G UACAAAAA	1604	UUUUUGUA GCcgaaagGCGaGuCaaGGuCu AAUAUGUU	9175
944	AUCAAAAU G UGUUUUAG	1605	CUAAAACA GCcgaaagGCGaGuCaaGGuCu AUUUUGAU	9176
946	CAAAAUGU G UUUUAGGA	1606	UCCUAAAA GCcgaaagGCGaGuCaaGGuCu ACAUUUUG	9177
963	AACUUCCU G UAAACAGG	1607	CCUGUUUA GCcgaaagGCGaGuCaaGGuCu AGGAAGUU	9178
991	GAAAGUAU G UCAACGAA	1608	UUCGUUGA GCcgaaagGCGaGuCaaGGuCu AUACUUUC	9179
1002	AACGAAUU G UGGGUCUU	1609	AAGACCCA GCcgaaagGCGaGuCaaGGuCu AAUUCGUU	9180
1039	CACGCAAU G UGGAUAUU	1610	AAUAUCCA GCcgaaagGCGaGuCaaGGuCu AUUGCGUG	9181
1137	AACAGUAU G UGAACCUU	1611	AAGGUUCA GCcgaaagGCGaGuCaaGGuCu AUACUGUU	9182
1184	UGCCAAGU G UUUGCUGA	1612	UCAGCAAA GCcgaaagGCGaGuCaaGGuCu ACUUGGCA	9183
1251	GAACCUUU G UGUCUCCU	1613	AGGAGACA GCcgaaagGCGaGuCaaGGuCu AAAGGUUC	9184
1253	ACCUUUGU G UCUCCUCU	1614	AGAGGAGA GCcgaaagGCGaGuCaaGGuCu ACAAAGGU	9185
1294	AGCCGCUU G UUUUGCUC	1615	GAGCAAAA GCcgaaagGCGaGuCaaGGuCu AAGCGGCU	9186
1344	ACAAUUCU G UCGUGCUC	1616	GAGCACGA GCcgaaagGCGaGuCaaGGuCu AGAAUUGU	9187
1390	GCUAGGCU G UGCUGCCA	1617	UGGCAGCA GCcgaaagGCGaGuCaaGGuCu AGCCUAGC	9188
1425	CGUCCUUU G UUUACGUC	1618	GACGUAAA GCcgaaagGCGaGuCaaGGuCu AAAGGACG	9189
1508	CGCCUAUU G UACCGACC	1619	GGUCGGUA GCcgaaagGCGaGuCaaGGuCu AAUAGGCG	9190
1557	CCCCGUCU G UGCCUUCU	1620	AGAAGGCA GCcgaaagGCGaGuCaaGGuCu AGACGGGG	9191
1581	CGGACCGU G UGCACUUC	1621	GAAGUGCA GCcgaaagGCGaGuCaaGGuCu ACGGUCCG	9192
1684	UCAGCAAU G UCAACGAC	1622	GUCGUUGA GCcgaaagGCGaGuCaaGGuCu AUUGCUGA	9193
1719	CAAAGACU G UGUGUUUA	1623	UAAACACA GCcgaaagGCGaGuCaaGGuCu AGUCUUUG	9194
1721	AAGACUGU G UGUUUAAU	1624	AUUAAACA GCcgaaagGCGaGuCaaGGuCu ACAGUCUU	9195
1723	GACUGUGU G UUUAAUGA	1625	UCAUUAAA GCcgaaagGCGaGuCaaGGuCu ACACAGUC	9196
1772	AGGUCUUU G UACUAGGA	1626	UCCUAGUA GCcgaaagGCGaGuCaaGGuCu AAAGACCU	9197
1785	AGGAGGCU G UAGGCAUA	1627	UAUGCCUA GCcgaaagGCGaGuCaaGGuCu AGCCUCCU	9198
1801	AAAUUGGU G UGUUCACC	1628	GGUGAACA GCcgaaagGCGaGuCaaGGuCu ACCAAUUU	9199
1803	AUUGGUGU G UUCACCAG	1629	CUGGUGAA GCcgaaagGCGaGuCaaGGuCu ACACCAAU	9200
1850	CAUCUCAU G UUCAUGUC	1630	GACAUGAA GCcgaaagGCGaGuCaaGGuCu AUGAGAUG	9201
1856	AUGUUCAU G UCCUACUG	1631	CAGUAGGA GCcgaaagGCGaGuCaaGGuCu AUGAACAU	9202
1864	GUCCUACU G UUCAAGCC	1632	GGCUUGAA GCcgaaagGCGaGuCaaGGuCu AGUAGGAC	9203
1881	UCCAAGCU G UGCCUUGG	1633	CCAAGGCA GCcgaaagGCGaGuCaaGGuCu AGCUUGGA	9204

1939	GAGCUUCU G UGGAGUUA	1634	UAACUCCA GCcgaaagGCGaGuCaaGGuCu AGAAGCUC	9205
2013	UCUGCUCU G UAUCGGGG	1635	CCCCGAUA GCcgaaagGCGaGuCaaGGuCu AGAGCAGA	9206
2045	GGAACAUU G UUCACCUC	1636	GAGGUGAA GCcgaaagGCGaGuCaaGGuCu AAUGUUCC	9207
2082	GCUAUUCU G UGUUGGGG	1637	CCCCAACA GCcgaaagGCGaGuCaaGGuCu AGAAUAGC	9208
2084	UAUUCUGU G UUGGGGUG	1638	CACCCCAA GCcgaaagGCGaGuCaaGGuCu ACAGAAUA	9209
2167	UCAGCUAU G UCAACGUU	1639	AACGUUGA GCcgaaagGCGaGuCaaGGuCu AUAGCUGA	9210
2205	CAACUAUU G UGGUUUCA	1640	UGAAACCA GCcgaaagGCGaGuCaaGGuCu AAUAGUUG	9211
2222	CAUUUCCU G UCUUACUU	1641	AAGUAAGA GCcgaaagGCGaGuCaaGGuCu AGGAAAUG	9212
2245	GAGAAACU G UUCUUGAA	1642	UUCAAGAA GCcgaaagGCGaGuCaaGGuCu AGUUUCUC	9213
2262	UAUUUGGU G UCUUUUGG	1643	CCAAAAGA GCcgaaagGCGaGuCaaGGuCu ACCAAAUA	9214
2274	UUUGGAGU G UGGAUUCG	1644	CGAAUCCA GCcgaaagGCGaGuCaaGGuCu ACUCCAAA	9215
2344	AAACUACU G UUGUUAGA	1645	UCUAACAA GCcgaaagGCGaGuCaaGGuCu AGUAGUUU	9216
2347	CUACUGUU G UUAGACGA	1646	UCGUCUAA GCcgaaagGCGaGuCaaGGuCu AACAGUAG	9217
2450	AUCUCAAU G UUAGUAUU	1647	AAUACUAA GCcgaaagGCGaGuCaaGGuCu AUUGAGAU	9218
2573	AGGACAUU G UUGAUAGA	1648	UCUAUCAA GCcgaaagGCGaGuCaaGGuCu AAUGUCCU	9219
2583	UGAUAGAU G UAAGCAAU	1649	AUUGCUUA GCcgaaagGCGaGuCaaGGuCu AUCUAUCA	9220
2594	AGCAAUUU G UGGGGCCC	1650	GGGCCCCA GCcgaaagGCGaGuCaaGGuCu AAAUUGCU	9221
2663	AUCCCAAU G UUACUAAA	1651	UUUAGUAA GCcgaaagGCGaGuCaaGGuCu AUUGGGAU	9222
2717	CAGAGUAU G UAGUUAAU	1652	AUUAACUA GCcgaaagGCGaGuCaaGGuCu AUACUCUG	9223
2901	AUCUUUCU G UCCCCAAU	1653	AUUGGGGA GCcgaaagGCGaGuCaaGGuCu AGAAAGAU	9224
3071	GGGGGACU G UUGGGGUG	1654	CACCCCAA GCcgaaagGCGaGuCaaGGuCu AGUCCCCC	9225
3111	UCACAACU G UGCCAGCA	1655	UGCUGGCA GCcgaaagGCGaGuCaaGGuCu AGUUGUGA	9226
40	AUCCCAGA G UCAGGGCC	1656	GGCCCUGA GCcgaaagGCGaGuCaaGGuCu UCUGGGAU	9227
46	GAGUCAGG G CCCUGUAC	1657	GUACAGGG GCcgaaagGCGaGuCaaGGuCu CCUGACUC	9228
65	UCCUGCUG G UGGCUCCA	1658	UGGAGCCA GCcgaaagGCGaGuCaaGGuCu CAGCAGGA	9229
68	UGCUGGUG G CUCCAGUU	1659	AACUGGAG GCcgaaagGCGaGuCaaGGuCu CACCAGCA	9230
74	UGGCUCCA G UUCAGGAA	1660	UUCCUGAA GCcgaaagGCGaGuCaaGGuCu UGGAGCCA	9231
85	CAGGAACA G UGAGCCCU	1661	AGGGCUCA GCcgaaagGCGaGuCaaGGuCu UGUUCCUG	9232
89	AACAGUGA G CCCUGCUC	1662	GAGCAGGG GCcgaaagGCGaGuCaaGGuCu UCACUGUU	9233
120	GCCAUAUC G UCAAUCUU	1663	AAGAUUGA GCcgaaagGCGaGuCaaGGuCu GAUAUGGC	9234
196	CCCUGCUC G UGUUACAG	1664	CUGUAACA GCcgaaagGCGaGuCaaGGuCu GAGCAGGG	9235
205	UGUUACAG G CGGGGUUU	1665	AAACCCCG GCcgaaagGCGaGuCaaGGuCu CUGUAACA	9236
210	CAGGCGGG G UUUUUCUU	1666	AAGAAAAA GCcgaaagGCGaGuCaaGGuCu CCCGCCUG	9237
248	ACCACAGA G UCUAGACU	1667	AGUCUAGA GCcgaaagGCGaGuCaaGGuCu UCUGUGGU	9238
258	CUAGACUC G UGGUGGAC	1668	GUCCACCA GCcgaaagGCGaGuCaaGGuCu GAGUCUAG	9239
261	GACUCGUG G UGGACUUC	1669	GAAGUCCA GCcgaaagGCGaGuCaaGGuCu CACGAGUC	9240
295	GAACACCC G UGUGUCUU	1670	AAGACACA GCcgaaagGCGaGuCaaGGuCu GGGUGUUC	9241
305	GUGUCUUG G CCAAAAUU	1671	AAUUUUGG GCcgaaagGCGaGuCaaGGuCu CAAGACAC	9242
318	AAUUCGCA G UCCCAAAU	1672	AUUUGGGA GCcgaaagGCGaGuCaaGGuCu UGCGAAUU	9243
332	AAUCUCCA G UCACUCAC	1673	GUGAGUGA GCcgaaagGCGaGuCaaGGuCu UGGAGAUU	9244
368	UUGUCCUG G UUAUCGCU	1674	AGCGAUAA GCcgaaagGCGaGuCaaGGuCu CAGGACAA	9245
390	UGUCUGCG G CGUUUUAU	1675	AUAAAACG GCcgaaagGCGaGuCaaGGuCu CGCAGACA	9246
392	UCUGCGGC G UUUUAUCA	1676	UGAUAAAA GCcgaaagGCGaGuCaaGGuCu GCCGCAGA	9247
442	UCUUGUUG G UUCUUCUG	1677	CAGAAGAA GCcgaaagGCGaGuCaaGGuCu CAACAAGA	9248
461	CUAUCAAG G UAUGUUGC	1678	GCAACAUA GCcgaaagGCGaGuCaaGGuCu CUUGAUAG	9249
472	UGUUGCCC G UUUGUCCU	1679	AGGACAAA GCcgaaagGCGaGuCaaGGuCu GGGCAACA	9250
506	AACAACCA G CACCGGAC	1680	GUCCGGUG GCcgaaagGCGaGuCaaGGuCu UGGUUGUU	9251
625	CAUCUUGG G CUUUCGCA	1681	UGCGAAAG GCcgaaagGCGaGuCaaGGuCu CCAAGAUG	9252
648	CUAUGGGA G UGGGCCUC	1682	GAGGCCCA GCcgaaagGCGaGuCaaGGuCu UCCCAUAG	9253
652	GGGAGUGG G CCUCAGUC	1683	GACUGAGG GCcgaaagGCGaGuCaaGGuCu CCACUCCC	9254
658	GGGCCUCA G UCCGUUUC	1684	GAAACGGA GCcgaaagGCGaGuCaaGGuCu UGAGGCCC	9255

662	CUCAGUCC G UUUCUCUU	1685	AAGAGAAA GCcgaaagGCGaGuCaaGGuCu GGACUGAG	9256
672	UUCUCUUG G CUCAGUUU	1686	AAACUGAG GCcgaaagGCGaGuCaaGGuCu CAAGAGAA	9257
677	UUGGCUCA G UUUACUAG	1687	CUAGUAAA GCcgaaagGCGaGuCaaGGuCu UGAGCCAA	9258
685	GUUUACUA G UGCCAUUU	1688	AAAUGGCA GCcgaaagGCGaGuCaaGGuCu UAGUAAAC	9259
699	UUUGUUCA G UGGUUCGU	1689	ACGAACCA GCcgaaagGCGaGuCaaGGuCu UGAACAAA	9260
702	GUUCAGUG G UUCGUAGG	1690	CCUACGAA GCcgaaagGCGaGuCaaGGuCu CACUGAAC	9261
706	AGUGGUUC G UAGGGCUU	1691	AAGCCCUA GCcgaaagGCGaGuCaaGGuCu GAACCACU	9262
711	UUCGUAGG G CUUUCCCC	1692	GGGGAAAG GCcgaaagGCGaGuCaaGGuCu CCUACGAA	9263
729	ACUGUCUG G CUUUCAGU	1693	ACUGAAAG GCcgaaagGCGaGuCaaGGuCu CAGACAGU	9264
736	GGCUUUCA G UUAUAUGG	1694	CCAUAUAA GCcgaaagGCGaGuCaaGGuCu UGAAAGCC	9265
753	AUGAUGUG G UUUUGGGG	1695	CCCCAAAA GCcgaaagGCGaGuCaaGGuCu CACAUCAU	9266
762	UUUUGGGG G CCAAGUCU	1696	AGACUUGG GCcgaaagGCGaGuCaaGGuCu CCCCAAAA	9267
767	GGGGCCAA G UCUGUACA	1697	UGUACAGA GCcgaaagGCGaGuCaaGGuCu UUGGCCCC	9268
785	CAUCUUGA G UCCCUUUA	1698	UAAAGGGA GCcgaaagGCGaGuCaaGGuCu UCAAGAUG	9269
826	GUCUUUGG G UAUACAUU	1699	AAUGUAUA GCcgaaagGCGaGuCaaGGuCu CCAAAGAC	9270
898	AAUUGGGA G UUGGGGCA	1700	UGCCCCAA GCcgaaagGCGaGuCaaGGuCu UCCCAAUU	9271
904	GAGUUGGG G CACAUUGC	1701	GCAAUGUG GCcgaaagGCGaGuCaaGGuCu CCCAACUC	9272
971	GUAAACAG G CCUAUUGA	1702	UCAAUAGG GCcgaaagGCGaGuCaaGGuCu CUGUUUAC	9273
987	AUUGGAAA G UAUGUCAA	1703	UUGACAUA GCcgaaagGCGaGuCaaGGuCu UUUCCAAU	9274
1006	AAUUGUGG G UCUUUUGG	1704	CCAAAAGA GCcgaaagGCGaGuCaaGGuCu CCACAAUU	9275
1016	CUUUUGGG G UUUGCCGC	1705	GCGGCAAA GCcgaaagGCGaGuCaaGGuCu CCCAAAAG	9276
1080	GCAUACAA G CAAAACAG	1706	CUGUUUUG GCcgaaagGCGaGuCaaGGuCu UUGUAUGC	9277
1089	CAAAACAG G CUUUUACU	1707	AGUAAAAG GCcgaaagGCGaGuCaaGGuCu CUGUUUUG	9278
1116	CUUACAAG G CCUUUCUA	1708	UAGAAAGG GCcgaaagGCGaGuCaaGGuCu CUUGUAAG	9279
1126	CUUUCUAA G UAAACAGU	1709	ACUGUUUA GCcgaaagGCGaGuCaaGGuCu UUAGAAAG	9280
1133	AGUAAACA G UAUGUGAA	1710	UUCACAUA GCcgaaagGCGaGuCaaGGuCu UGUUUACU	9281
1152	UUUACCCC G UUGCUCGG	1711	CCGAGCAA GCcgaaagGCGaGuCaaGGuCu GGGGUAAA	9282
1160	GUUGCUCG G CAACGGCC	1712	GGCCGUUG GCcgaaagGCGaGuCaaGGuCu CGAGCAAC	9283
1166	CGGCAACG G CCUGGUCU	1713	AGACCAGG GCcgaaagGCGaGuCaaGGuCu CGUUGCCG	9284
1171	ACGGCCUG G UCUAUGCC	1714	GGCAUAGA GCcgaaagGCGaGuCaaGGuCu CAGGCCGU	9285
1182	UAUGCCAA G UGUUUGCU	1715	AGCAAACA GCcgaaagGCGaGuCaaGGuCu UUGGCAUA	9286
1207	CCCCACUG G UUGGGGCU	1716	AGCCCCAA GCcgaaagGCGaGuCaaGGuCu CAGUGGGG	9287
1213	UGGUUGGG G CUUGGCCA	1717	UGGCCAAG GCcgaaagGCGaGuCaaGGuCu CCCAACCA	9288
1218	GGGGCUUG G CCAUAGGC	1718	GCCUAUGG GCcgaaagGCGaGuCaaGGuCu CAAGCCCC	9289
1225	GGCCAUAG G CCAUCAGC	1719	GCUGAUGG GCcgaaagGCGaGuCaaGGuCu CUAUGGCC	9290
1232	GGCCAUCA G CGCAUGCG	1720	CGCAUGCG GCcgaaagGCGaGuCaaGGuCu UGAUGGCC	9291
1240	GCGCAUGC G UGGAACCU	1721	AGGUUCCA GCcgaaagGCGaGuCaaGGuCu GCAUGCGC	9292
1287	AACUCCUA G CCGCUUGU	1722	ACAAGCGG GCcgaaagGCGaGuCaaGGuCu UAGGAGUU	9293
1306	UGCUCGCA G CAGGUCUG	1723	CAGACCUG GCcgaaagGCGaGuCaaGGuCu UGCGAGCA	9294
1310	CGCAGCAG G UCUGGGGC	1724	GCCCCAGA GCcgaaagGCGaGuCaaGGuCu CUGCUGCG	9295
1317	GGUCUGGG G CAAAACUC	1725	GAGUUUUG GCcgaaagGCGaGuCaaGGuCu CCCAGACC	9296
1347	AUUCUGUC G UGCUCUCC	1726	GGAGAGCA GCcgaaagGCGaGuCaaGGuCu GACAGAAU	9297
1379	UUUCCAUG G CUGCUAGG	1727	CCUAGCAG GCcgaaagGCGaGuCaaGGuCu CAUGGAAA	9298
1387	GCUGCUAG G CUGUGCUG	1728	CAGCACAG GCcgaaagGCGaGuCaaGGuCu CUAGCAGC	9299
1418	CGCGGGAC G UCCUUUGU	1729	ACAAAGGA GCcgaaagGCGaGuCaaGGuCu GUCCCGCG	9300
1431	UUGUUUAC G UCCCGUCG	1730	CGACGGGA GCcgaaagGCGaGuCaaGGuCu GUAAACAA	9301
1436	UACGUCCC G UCGGCGCU	1731	AGCGCCGA GCcgaaagGCGaGuCaaGGuCu GGGACGUA	9302
1440	UCCCGUCG G CGCUGAAU	1732	AUUCAGCG GCcgaaagGCGaGuCaaGGuCu CGACGGGA	9303
1471	CUCCCGGG G CCGCUUGG	1733	CCAAGCGG GCcgaaagGCGaGuCaaGGuCu CCCGGGAG	9304
1481	CGCUUGGG G CUCUACCG	1734	CGGUAGAG GCcgaaagGCGaGuCaaGGuCu CCCAAGCG	9305
1517	UACCGACC G UCCACGGG	1735	CCCGUGGA GCcgaaagGCGaGuCaaGGuCu GGUCGGUA	9306

1526	UCCACGGG G CGCACCUC	1736	GAGGUGCG GCcgaaagGCGaGuCaaGGuCu CCCGUGGA	9307
1553	GACUCCCC G UCUGUGCC	1737	GGCACAGA GCcgaaagGCGaGuCaaGGuCu GGGGAGUC	9308
1579	GCCGGACC G UGUGCACU	1738	AGUGCACA GCcgaaagGCGaGuCaaGGuCu GGUCCGGC	9309
1605	CUCUGCAC G UCGCAUGG	1739	CCAUGCGA GCcgaaagGCGaGuCaaGGuCu GUGCAGAG	9310
1622	AGACCACC G UGAACGCC	1740	GGCGUUCA GCcgaaagGCGaGuCaaGGuCu GGUGGUCU	9311
1649	UGCCCAAG G UCUUGCAU	1741	AUGCAAGA GCcgaaagGCGaGuCaaGGuCu CUUGGGCA	9312
1679	GACUUUCA G CAAUGUCA	1742	UGACAUUG GCcgaaagGCGaGuCaaGGuCu UGAAAGUC	9313
1703	ACCUUGAG G CAUACUUC	1743	GAAGUAUG GCcgaaagGCGaGuCaaGGuCu CUCAAGGU	9314
1732	UUUAAUGA G UGGGAGGA	1744	UCCUCCCA GCcgaaagGCGaGuCaaGGuCu UCAUUAAA	9315
1741	UGGGAGGA G UUGGGGGA	1745	UCCCCCAA GCcgaaagGCGaGuCaaGGuCu UCCUCCCA	9316
1754	GGGAGGAG G UUAGGUUA	1746	UAACCUAA GCcgaaagGCGaGuCaaGGuCu CUCCUCCC	9317
1759	GAGGUUAG G UUAAAGGU	1747	ACCUUUAA GCcgaaagGCGaGuCaaGGuCu CUAACCUC	9318
1766	GGUUAAAG G UCUUUGUA	1748	UACAAAGA GCcgaaagGCGaGuCaaGGuCu CUUUAACC	9319
1782	ACUAGGAG G CUGUAGGC	1749	GCCUACAG GCcgaaagGCGaGuCaaGGuCu CUCCUAGU	9320
1789	GGCUGUAG G CAUAAAUU	1750	AAUUUAUG GCcgaaagGCGaGuCaaGGuCu CUACAGCC	9321
1799	AUAAAUUG G UGUGUUCA	1751	UGAACACA GCcgaaagGCGaGuCaaGGuCu CAAUUUAU	9322
1811	GUUCACCA G CACCAUGC	1752	GCAUGGUG GCcgaaagGCGaGuCaaGGuCu UGGUGAAC	9323
1870	CUGUUCAA G CCUCCAAG	1753	CUUGGAGG GCcgaaagGCGaGuCaaGGuCu UUGAACAG	9324
1878	GCCUCCAA G CUGUGCCU	1754	AGGCACAG GCcgaaagGCGaGuCaaGGuCu UUGGAGGC	9325
1890	UGCCUUGG G UGGCUUUG	1755	CAAAGCCA GCcgaaagGCGaGuCaaGGuCu CCAAGGCA	9326
1893	CUUGGGUG G CUUUGGGG	1756	CCCCAAAG GCcgaaagGCGaGuCaaGGuCu CACCCAAG	9327
1901	GCUUUGGG G CAUGGACA	1757	UGUCCAUG GCcgaaagGCGaGuCaaGGuCu CCCAAAGC	9328
1917	AUUGACCC G UAUAAAGA	1758	UCUUUAUA GCcgaaagGCGaGuCaaGGuCu GGGUCAAU	9329
1933	AAUUUGGA G CUUCUGUG	1759	CACAGAAG GCcgaaagGCGaGuCaaGGuCu UCCAAAUU	9330
1944	UCUGUGGA G UUACUCUC	1760	GAGAGUAA GCcgaaagGCGaGuCaaGGuCu UCCACAGA	9331
2023	AUCGGGGG G CCUUAGAG	1761	CUCUAAGG GCcgaaagGCGaGuCaaGGuCu CCCCCGAU	9332
2031	GCCUUAGA G UCUCCGGA	1762	UCCGGAGA GCcgaaagGCGaGuCaaGGuCu UCUAAGGC	9333
2062	ACCAUACG G CACUCAGG	1763	CCUGAGUG GCcgaaagGCGaGuCaaGGuCu CGUAUGGU	9334
2070	GCACUCAG G CAAGCUAU	1764	AUAGCUUG GCcgaaagGCGaGuCaaGGuCu CUGAGUGC	9335
2074	UCAGGCAA G CUAUUCUG	1765	CAGAAUAG GCcgaaagGCGaGuCaaGGuCu UUGCCUGA	9336
2090	GUGUUGGG G UGAGUUGA	1766	UCAACUCA GCcgaaagGCGaGuCaaGGuCu CCCAACAC	9337
2094	UGGGGUGA G UUGAUGAA	1767	UUCAUCAA GCcgaaagGCGaGuCaaGGuCu UCACCCCA	9338
2107	UGAAUCUA G CCACCUGG	1768	CCAGGUGG GCcgaaagGCGaGuCaaGGuCu UAGAUUCA	9339
2116	CCACCUGG G UGGGAAGU	1769	ACUUCCCA GCcgaaagGCGaGuCaaGGuCu CCAGGUGG	9340
2123	GGUGGGAA G UAAUUUGG	1770	CCAAAUUA GCcgaaagGCGaGuCaaGGuCu UUCCCACC	9341
2140	AAGAUCCA G CAUCCAGG	1771	CCUGGAUG GCcgaaagGCGaGuCaaGGuCu UGGAUCUU	9342
2155	GGGAAUUA G UAGUCAGC	1772	GCUGACUA GCcgaaagGCGaGuCaaGGuCu UAAUUCCC	9343
2158	AAUUAGUA G UCAGCUAU	1773	AUAGCUGA GCcgaaagGCGaGuCaaGGuCu UACUAAUU	9344
2162	AGUAGUCA G CUAUGUCA	1774	UGACAUAG GCcgaaagGCGaGuCaaGGuCu UGACUACU	9345
2173	AUGUCAAC G UUAAUAUG	1775	CAUAUUAA GCcgaaagGCGaGuCaaGGuCu GUUGACAU	9346
2183	UAAUAUGG G CCUAAAAA	1776	UUUUUAGG GCcgaaagGCGaGuCaaGGuCu CCAUAUUA	9347
2208	CUAUUGUG G UUUCACAU	1777	AUGUGAAA GCcgaaagGCGaGuCaaGGuCu CACAAUAG	9348
2235	ACUUUUGG G CGAGAAAC	1778	GUUUCUCG GCcgaaagGCGaGuCaaGGuCu CCAAAAGU	9349
2260	AAUAUUUG G UGUCUUUU	1779	AAAAGACA GCcgaaagGCGaGuCaaGGuCu CAAAUAUU	9350
2272	CUUUUGGA G UGUGGAUU	1780	AAUCCACA GCcgaaagGCGaGuCaaGGuCu UCCAAAAG	9351
2360	ACGAAGAG G CAGGUCCC	1781	GGGACCUG GCcgaaagGCGaGuCaaGGuCu CUCUUCGU	9352
2364	AGAGGCAG G UCCCCUAG	1782	CUAGGGGA GCcgaaagGCGaGuCaaGGuCu CUGCCUCU	9353
2403	AGACGAAG G UCUCAAUC	1783	GAUUGAGA GCcgaaagGCGaGuCaaGGuCu CUUCGUCU	9354
2417	AUCGCCGC G UCGCAGAA	1784	UUCUGCGA GCcgaaagGCGaGuCaaGGuCu GCGGCGAU	9355
2454	CAAUGUUA G UAUUCCUU	1785	AAGGAAUA GCcgaaagGCGaGuCaaGGuCu UAACAUUG	9356
2474	CACAUAAG G UGGGAAAC	1786	GUUUCCCA GCcgaaagGCGaGuCaaGGuCu CUUAUGUG	9357

2491	UUUACGGG G CUUUAUUC	1787	GAAUAAAG GCcgaaagGCGaGuCaaGGuCu CCCGUAAA	9358
2507	CUUCUACG G UACCUUGC	1788	GCAAGGUA GCcgaaagGCGaGuCaaGGuCu CGUAGAAG	9359
2530	CCUAAAUG G CAAACUCC	1789	GGAGUUUG GCcgaaagGCGaGuCaaGGuCu CAUUUAGG	9360
2587	AGAUGUAA G CAAUUUGU	1790	ACAAAUUG GCcgaaagGCGaGuCaaGGuCu UUACAUCU	9361
2599	UUUGUGGG G CCCCUUAC	1791	GUAAGGGG GCcgaaagGCGaGuCaaGGuCu CCCACAAA	9362
2609	CCCUUACA G UAAAUGAA	1792	UUCAUUUA GCcgaaagGCGaGuCaaGGuCu UGUAAGGG	9363
2650	CCUGCUAG G UUUUAUCC	1793	GGAUAAAA GCcgaaagGCGaGuCaaGGuCu CUAGCAGG	9364
2701	AUCAAACC G UAUUAUCC	1794	GGAUAAUA GCcgaaagGCGaGuCaaGGuCu GGUUUGAU	9365
2713	UAUCCAGA G UAUGUAGU	1795	ACUACAUA GCcgaaagGCGaGuCaaGGuCu UCUGGAUA	9366
2720	AGUAUGUA G UUAAUCAU	1796	AUGAUUAA GCcgaaagGCGaGuCaaGGuCu UACAUACU	9367
2768	UUUGGAAG G CGGGGAUC	1797	GAUCCCCG GCcgaaagGCGaGuCaaGGuCu CUUCCAAA	9368
2791	AAAAGAGA G UCCACACG	1798	CGUGUGGA GCcgaaagGCGaGuCaaGGuCu UCUCUUUU	9369
2799	GUCCACAC G UAGCGCCU	1799	AGGCGCUA GCcgaaagGCGaGuCaaGGuCu GUGUGGAC	9370
2802	CACACGUA G CGCCUCAU	1800	AUGAGGCG GCcgaaagGCGaGuCaaGGuCu UACGUGUG	9371
2818	UUUUGCGG G UCACCAUA	1801	UAUGGUGA GCcgaaagGCGaGuCaaGGuCu CCGCAAAA	9372
2848	GAUCUACA G CAUGGGAG	1802	CUCCCAUG GCcgaaagGCGaGuCaaGGuCu UGUAGAUC	9373
2857	CAUGGGAG G UUGGUCUU	1803	AAGACCAA GCcgaaagGCGaGuCaaGGuCu CUCCCAUG	9374
2861	GGAGGUUG G UCUUCCAA	1804	UUGGAAGA GCcgaaagGCGaGuCaaGGuCu CAACCUCC	9375
2881	UCGAAAAG G CAUGGGGA	1805	UCCCCAUG GCcgaaagGCGaGuCaaGGuCu CUUUUCGA	9376
2936	GAUCAUCA G UUGGACCC	1806	GGGUCCAA GCcgaaagGCGaGuCaaGGuCu UGAUGAUC	9377
2955	CAUUCAAA G CCAACUCA	1807	UGAGUUGG GCcgaaagGCGaGuCaaGGuCu UUUGAAUG	9378
2964	CCAACUCA G UAAAUCCA	1808	UGGAUUUA GCcgaaagGCGaGuCaaGGuCu UGAGUUGG	9379
3005	GACAACUG G CCGGACGC	1809	GCGUCCGG GCcgaaagGCGaGuCaaGGuCu CAGUUGUC	9380
3021	CCAACAAG G UGGGAGUG	1810	CACUCCCA GCcgaaagGCGaGuCaaGGuCu CUUGUUGG	9381
3027	AGGUGGGA G UGGGAGCA	1811	UGCUCCCA GCcgaaagGCGaGuCaaGGuCu UCCCACCU	9382
3033	GAGUGGGA G CAUUCGGG	1812	CCCGAAUG GCcgaaagGCGaGuCaaGGuCu UCCCACUC	9383
3041	GCAUUCGG G CCAGGGUU	1813	AACCCUGG GCcgaaagGCGaGuCaaGGuCu CCGAAUGC	9384
3047	GGGCCAGG G UUCACCCC	1814	GGGGUGAA GCcgaaagGCGaGuCaaGGuCu CCUGGCCC	9385
3077	CUGUUGGG G UGGAGCCC	1815	GGGCUCCA GCcgaaagGCGaGuCaaGGuCu CCCAACAG	9386
3082	GGGGUGGA G CCCUCACG	1816	CGUGAGGG GCcgaaagGCGaGuCaaGGuCu UCCACCCC	9387
3097	CGCUCAGG G CCUACUCA	1817	UGAGUAGG GCcgaaagGCGaGuCaaGGuCu CCUGAGCG	9388
3117	CUGUGCCA G CAGCUCCU	1818	AGGAGCUG GCcgaaagGCGaGuCaaGGuCu UGGCACAG	9389
3120	UGCCAGCA G CUCCUCCU	1819	AGGAGGAG GCcgaaagGCGaGuCaaGGuCu UGCUGGCA	9390
3146	ACCAAUCG G CAGUCAGG	1820	CCUGACUG GCcgaaagGCGaGuCaaGGuCu CGAUUGGU	9391
3149	AAUCGGCA G UCAGGAAG	1821	CUUCCUGA GCcgaaagGCGaGuCaaGGuCu UGCCGAUU	9392
3158	UCAGGAAG G CAGCCUAC	1822	GUAGGCUG GCcgaaagGCGaGuCaaGGuCu CUUCCUGA	9393
3161	GGAAGGCA G CCUACUCC	1823	GGAGUAGG GCcgaaagGCGaGuCaaGGuCu UGCCUUCC	9394
3204	AUCCUCAG G CCAUGCAG	1824	CUGCAUGG GCcgaaagGCGaGuCaaGGuCu CUGAGGAU	9395

Input Sequence = AF100308. Cut Site = YG/M or UG/U.
Stem Length = 8 . Core Sequence = GCcgaaagGCGaGuCaaGGuCu
AF100308 (Hepatitis B virus strain 2-18, 3215 bp)

TABLE IX: HUMAN HBV DNAZYME AND SUBSTRATE SEQUENCE

Pos	Substrate	Seq ID	DNAzyme	Seq ID
508	CAACCAGC A CCGGACCA	833	TGGTCCGG GGCTAGCTACAACGA GCTGGTTG	9396
1632	GAACGCCC A CAGGAACC	1096	GGTTCCTG GGCTAGCTACAACGA GGGCGTTC	9397
2992	CAACCCGC A CAAGGACA	1376	TGTCCTTG GGCTAGCTACAACGA GCGGGTTG	9398
61	ACUUUCCU G CUGGUGGC	1448	GCCACCAG GGCTAGCTACAACGA AGGAAAGT	9399
94	UGAGCCCU G CUCAGAAU	1450	ATTCTGAG GGCTAGCTACAACGA AGGGCTCA	9400
112	CUGUCUCU G CCAUAUCG	1451	CGATATGG GGCTAGCTACAACGA AGAGACAG	9401
169	AGAACAUC G CAUCAGGA	1454	TCCTGATG GGCTAGCTACAACGA GATGTTCT	9402
192	GGACCCCU G CUCGUGUU	1455	AACACGAG GGCTAGCTACAACGA AGGGGTCC	9403
315	CAAAAUUC G CAGUCCCA	1457	TGGGACTG GGCTAGCTACAACGA GAATTTTG	9404
374	UGGUUAUC G CUGGAUGU	1458	ACATCCAG GGCTAGCTACAACGA GATAACCA	9405
387	AUGUGUCU G CGGCGUUU	1459	AAACGCCG GGCTAGCTACAACGA AGACACAT	9406
410	CUUCCUCU G CAUCCUGC	1460	GCAGGATG GGCTAGCTACAACGA AGAGGAAG	9407
417	UGCAUCCU G CUGCUAUG	1461	CATAGCAG GGCTAGCTACAACGA AGGATGCA	9408
420	AUCCUGCU G CUAUGCCU	1462	AGGCATAG GGCTAGCTACAACGA AGCAGGAT	9409
425	GCUGCUAU G CCUCAUCU	1463	AGATGAGG GGCTAGCTACAACGA ATAGCAGC	9410
468	GGUAUGUU G CCCGUUUG	1464	CAAACGGG GGCTAGCTACAACGA AACATACC	9411
518	CGGACCAU G CAAAACCU	1465	AGGTTTTG GGCTAGCTACAACGA ATGGTCCG	9412
527	CAAAACCU G CACAACUC	1466	GAGTTGTG GGCTAGCTACAACGA AGGTTTTG	9413
538	CAACUCCU G CUCAAGGA	1467	TCCTTGAG GGCTAGCTACAACGA AGGAGTTG	9414
569	CUCAUGUU G CUGUACAA	1468	TTGTACAG GGCTAGCTACAACGA AACATGAG	9415
596	CGGAAACU G CACCUGUA	1469	TACAGGTG GGCTAGCTACAACGA AGTTTCCG	9416
631	GGGCUUUC G CAAAAUAC	1470	GTATTTTG GGCTAGCTACAACGA GAAAGCCC	9417
687	UUACUAGU G CCAUUUGU	1471	ACAAATGG GGCTAGCTACAACGA ACTAGTAA	9418
795	CCCUUUAU G CCGCUGUU	1474	AACAGCGG GGCTAGCTACAACGA ATAAAGGG	9419
798	UUUAUGCC G CUGUUACC	1475	GGTAACAG GGCTAGCTACAACGA GGCATAAA	9420
911	GGCACAUU G CCACAGGA	1476	TCCTGTGG GGCTAGCTACAACGA AATGTGCC	9421
1020	UGGGGUUU G CCGCCCCU	1479	AGGGGCGG GGCTAGCTACAACGA AAACCCCA	9422
1023	GGUUUGCC G CCCCUUUC	1480	GAAAGGGG GGCTAGCTACAACGA GGCAAACC	9423
1034	CCUUUCAC G CAAUGUGG	1481	CCACATTG GGCTAGCTACAACGA GTGAAAGG	9424
1050	GAUAUUCU G CUUUAAUG	1482	CATTAAAG GGCTAGCTACAACGA AGAATATC	9425
1058	GCUUUAAU G CCUUUAUA	1483	TATAAAGG GGCTAGCTACAACGA ATTAAAGC	9426
1068	CUUUAUAU G CAUGCAUA	1484	TATGCATG GGCTAGCTACAACGA ATATAAAG	9427
1072	AUAUGCAU G CAUACAAG	1485	CTTGTATG GGCTAGCTACAACGA ATGCATAT	9428
1103	ACUUUCUC G CCAACUUA	1486	TAAGTTGG GGCTAGCTACAACGA GAGAAAGT	9429
1155	ACCCCGUU G CUCGGCAA	1488	TTGCCGAG GGCTAGCTACAACGA AACGGGGT	9430
1177	UGGUCUAU G CCAAGUGU	1489	ACACTTGG GGCTAGCTACAACGA ATAGACCA	9431
1188	AAGUGUUU G CUGACGCA	1490	TGCGTCAG GGCTAGCTACAACGA AAACACTT	9432
1194	UUGCUGAC G CAACCCCC	1492	GGGGGTTG GGCTAGCTACAACGA GTCAGCAA	9433
1234	CCAUCAGC G CAUGCGUG	1493	CACGCATG GGCTAGCTACAACGA GCTGATGG	9434
1238	CAGCGCAU G CGUGGAAC	1494	GTTCCACG GGCTAGCTACAACGA ATGCGCTG	9435
1262	UCUCCUCU G CCGAUCCA	1495	TGGATCGG GGCTAGCTACAACGA AGAGGAGA	9436
1275	UCCAUACC G CGGAACUC	1497	GAGTTCCG GGCTAGCTACAACGA GGTATGGA	9437
1290	UCCUAGCC G CUUGUUUU	1498	AAAACAAG GGCTAGCTACAACGA GGCTAGGA	9438
1299	CUUGUUUU G CUCGCAGC	1499	GCTGCGAG GGCTAGCTACAACGA AAAACAAG	9439
1303	UUUUGCUC G CAGCAGGU	1500	ACCTGCTG GGCTAGCTACAACGA GAGCAAAA	9440
1349	UCUGUCGU G CUCUCCCG	1502	CGGGAGAG GGCTAGCTACAACGA ACGACAGA	9441
1357	GCUCUCCC G CAAAUAUA	1503	TATATTTG GGCTAGCTACAACGA GGGAGAGC	9442

1382	CCAUGGCU G CUAGGCUG	1504	CAGCCTAG GGCTAGCTACAACGA AGCCATGG	9443
1392	UAGGCUGU G CUGCCAAC	1505	GTTGGCAG GGCTAGCTACAACGA ACAGCCTA	9444
1395	GCUGUGCU G CCAACUGG	1506	CCAGTTGG GGCTAGCTACAACGA AGCACAGC	9445
1411	GAUCCUAC G CGGGACGU	1507	ACGTCCCG GGCTAGCTACAACGA GTAGGATC	9446
1442	CCGUCGGC G CUGAAUCC	1508	GGATTCAG GGCTAGCTACAACGA GCCGACGG	9447
1452	UGAAUCCC G CGGACGAC	1510	GTCGTCCG GGCTAGCTACAACGA GGGATTCA	9448
1474	CCGGGGCC G CUUGGGGC	1512	GCCCCAAG GGCTAGCTACAACGA GGCCCCGG	9449
1489	GCUCUACC G CCCGCUUC	1513	GAAGCGGG GGCTAGCTACAACGA GGTAGAGC	9450
1493	UACCGCCC G CUUCUCCG	1514	CGGAGAAG GGCTAGCTACAACGA GGGCGGTA	9451
1501	GCUUCUCC G CCUAUUGU	1515	ACAATAGG GGCTAGCTACAACGA GGAGAAGC	9452
1528	CACGGGGC G CACCUCUC	1517	GAGAGGTG GGCTAGCTACAACGA GCCCCGTG	9453
1542	CUCUUUAC G CGGACUCC	1518	GGAGTCCG GGCTAGCTACAACGA GTAAAGAG	9454
1559	CCGUCUGU G CCUUCUCA	1519	TGAGAAGG GGCTAGCTACAACGA ACAGACGG	9455
1571	UCUCAUCU G CCGGACCG	1520	CGGTCCGG GGCTAGCTACAACGA AGATGAGA	9456
1583	GACCGUGU G CACUUCGC	1521	GCGAAGTG GGCTAGCTACAACGA ACACGGTC	9457
1590	UGCACUUC G CUUCACCU	1522	AGGTGAAG GGCTAGCTACAACGA GAAGTGCA	9458
1601	UCACCUCU G CACGUCGC	1523	GCGACGTG GGCTAGCTACAACGA AGAGGTGA	9459
1608	UGCACGUC G CAUGGAGA	1524	TCTCCATG GGCTAGCTACAACGA GACGTGCA	9460
1628	CCGUGAAC G CCCACAGG	1526	CCTGTGGG GGCTAGCTACAACGA GTTCACGG	9461
1642	AGGAACCU G CCCAAGGU	1527	ACCTTGGG GGCTAGCTACAACGA AGGTTCCT	9462
1654	AAGGUCUU G CAUAAGAG	1528	CTCTTATG GGCTAGCTACAACGA AAGACCTT	9463
1818	AGCACCAU G CAACUUUU	1533	AAAAGTTG GGCTAGCTACAACGA ATGGTGCT	9464
1835	UCACCUCU G CCUAAUCA	1534	TGATTAGG GGCTAGCTACAACGA AGAGGTGA	9465
1883	CAAGCUGU G CCUUGGGU	1535	ACCCAAGG GGCTAGCTACAACGA ACAGCTTG	9466
1959	UCUUUUUU G CCUUCUGA	1537	TCAGAAGG GGCTAGCTACAACGA AAAAAAGA	9467
2002	UCGACACC G CCUCUGCU	1541	AGCAGAGG GGCTAGCTACAACGA GGTGTCGA	9468
2008	CCGCCUCU G CUCUGUAU	1542	ATACAGAG GGCTAGCTACAACGA AGAGGCGG	9469
2282	GUGGAUUC G CACUCCUC	1548	GAGGAGTG GGCTAGCTACAACGA GAATCCAC	9470
2293	CUCCUCCU G CAUAUAGA	1549	TCTATATG GGCTAGCTACAACGA AGGAGGAG	9471
2311	CACCAAAU G CCCCUAUC	1550	GATAGGGG GGCTAGCTACAACGA ATTTGGTG	9472
2388	ACUCCCUC G CCUCGCAG	1552	CTGCGAGG GGCTAGCTACAACGA GAGGGAGT	9473
2393	CUCGCCUC G CAGACGAA	1553	TTCGTCTG GGCTAGCTACAACGA GAGGCGAG	9474
2412	UCUCAAUC G CCGCGUCG	1555	CGACGCGG GGCTAGCTACAACGA GATTGAGA	9475
2415	CAAUCGCC G CGUCGCAG	1556	CTGCGACG GGCTAGCTACAACGA GGCGATTG	9476
2420	GCCGCGUC G CAGAAGAU	1557	ATCTTCTG GGCTAGCTACAACGA GACGCGGC	9477
2514	GGUACCUU G CUUUAAUC	1558	GATTAAAG GGCTAGCTACAACGA AAGGTACC	9478
2560	AUUCAUUU G CAGGAGGA	1560	TCCTCCTG GGCTAGCTACAACGA AAATGAAT	9479
2641	UUAACUAU G CCUGCUAG	1563	CTAGCAGG GGCTAGCTACAACGA ATAGTTAA	9480
2645	CUAUGCCU G CUAGGUUU	1564	AAACCTAG GGCTAGCTACAACGA AGGCATAG	9481
2677	AAAUAUUU G CCCUUAGA	1565	TCTAAGGG GGCTAGCTACAACGA AAATATTT	9482
2740	UUCCAGAC G CGACAUUA	1566	TAATGTCG GGCTAGCTACAACGA GTCTGGAA	9483
2804	CACGUAGC G CCUCAUUU	1568	AAATGAGG GGCTAGCTACAACGA GCTACGTG	9484
2814	CUCAUUUU G CGGGUCAC	1569	GTGACCCG GGCTAGCTACAACGA AAAATGAG	9485
2946	UGGACCCU G CAUUCAAA	1572	TTTGAATG GGCTAGCTACAACGA AGGGTCCA	9486
2990	CUCAACCC G CACAAGGA	1573	TCCTTGTG GGCTAGCTACAACGA GGGTTGAG	9487
3012	GGCCGGAC G CCAACAAG	1574	CTTGTTGG GGCTAGCTACAACGA GTCCGGCC	9488
3090	GCCCUCAC G CUCAGGGC	1575	GCCCTGAG GGCTAGCTACAACGA GTGAGGGC	9489
3113	ACAACUGU G CCAGCAGC	1576	GCTGCTGG GGCTAGCTACAACGA ACAGTTGT	9490
3132	CUCCUCCU G CCUCCACC	1577	GGTGGAGG GGCTAGCTACAACGA AGGAGGAG	9491
51	AGGGCCCU G UACUUUCC	1578	GGAAAGTA GGCTAGCTACAACGA AGGGCCCT	9492
106	AGAAUACU G UCUCUGCC	1579	GGCAGAGA GGCTAGCTACAACGA AGTATTCT	9493

148	GGGACCCU G UACCGAAC	1580	GTTCGGTA GGCTAGCTACAACGA AGGGTCCC	9494
198	CUGCUCGU G UUACAGGC	1581	GCCTGTAA GGCTAGCTACAACGA ACGAGCAG	9495
219	UUUUUCUU G UUGACAAA	1582	TTTGTCAA GGCTAGCTACAACGA AAGAAAAA	9496
297	ACACCCGU G UGUCUUGG	1583	CCAAGACA GGCTAGCTACAACGA ACGGGTGT	9497
299	ACCCGUGU G UCUUGGCC	1584	GGCCAAGA GGCTAGCTACAACGA ACACGGGT	9498
347	ACCAACCU G UUGUCCUC	1585	GAGGACAA GGCTAGCTACAACGA AGGTTGGT	9499
350	AACCUGUU G UCCUCCAA	1586	TTGGAGGA GGCTAGCTACAACGA AACAGGTT	9500
362	UCCAAUUU G UCCUGGUU	1587	AACCAGGA GGCTAGCTACAACGA AAATTGGA	9501
381	CGCUGGAU G UGUCUGCG	1588	CGCAGACA GGCTAGCTACAACGA ATCCAGCG	9502
383	CUGGAUGU G UCUGCGGC	1589	GCCGCAGA GGCTAGCTACAACGA ACATCCAG	9503
438	AUCUUCUU G UUGGUUCU	1590	AGAACCAA GGCTAGCTACAACGA AAGAAGAT	9504
465	CAAGGUAU G UUGCCCGU	1591	ACGGGCAA GGCTAGCTACAACGA ATACCTTG	9505
476	GCCCGUUU G UCCUCUAA	1592	TTAGAGGA GGCTAGCTACAACGA AAACGGGC	9506
555	ACCUCUAU G UUUCCCUC	1593	GAGGGAAA GGCTAGCTACAACGA ATAGAGGT	9507
566	UCCCUCAU G UUGCUGUA	1594	TACAGCAA GGCTAGCTACAACGA ATGAGGGA	9508
572	AUGUUGCU G UACAAAAC	1595	GTTTTGTA GGCTAGCTACAACGA AGCAACAT	9509
602	CUGCACCU G UAUUCCCA	1596	TGGGAATA GGCTAGCTACAACGA AGGTGCAG	9510
694	UGCCAUUU G UUCAGUGG	1597	CCACTGAA GGCTAGCTACAACGA AAATGGCA	9511
724	CCCCCACU G UCUGGCUU	1598	AAGCCAGA GGCTAGCTACAACGA AGTGGGGG	9512
750	UGGAUGAU G UGGUUUUG	1599	CAAAACCA GGCTAGCTACAACGA ATCATCCA	9513
771	CCAAGUCU G UACAACAU	1600	ATGTTGTA GGCTAGCTACAACGA AGACTTGG	9514
801	AUGCCGCU G UUACCAAU	1601	ATTGGTAA GGCTAGCTACAACGA AGCGGCAT	9515
818	UUUCUUUU G UCUUUGGG	1602	CCCAAAGA GGCTAGCTACAACGA AAAAGAAA	9516
888	UGGGAUAU G UAAUUGGG	1603	CCCAATTA GGCTAGCTACAACGA ATATCCCA	9517
927	AACAUAUU G UACAAAAA	1604	TTTTTGTA GGCTAGCTACAACGA AATATGTT	9518
944	AUCAAAAU G UGUUUUAG	1605	CTAAAACA GGCTAGCTACAACGA ATTTTGAT	9519
946	CAAAAUGU G UUUUAGGA	1606	TCCTAAAA GGCTAGCTACAACGA ACATTTTG	9520
963	AACUUCCU G UAAACAGG	1607	CCTGTTTA GGCTAGCTACAACGA AGGAAGTT	9521
991	GAAAGUAU G UCAACGAA	1608	TTCGTTGA GGCTAGCTACAACGA ATACTTTC	9522
1002	AACGAAUU G UGGGUCUU	1609	AAGACCCA GGCTAGCTACAACGA AATTCGTT	9523
1039	CACGCAAU G UGGAUAUU	1610	AATATCCA GGCTAGCTACAACGA ATTGCGTG	9524
1137	AACAGUAU G UGAACCUU	1611	AAGGTTCA GGCTAGCTACAACGA ATACTGTT	9525
1184	UGCCAAGU G UUUGCUGA	1612	TCAGCAAA GGCTAGCTACAACGA ACTTGGCA	9526
1251	GAACCUUU G UGUCUCCU	1613	AGGAGACA GGCTAGCTACAACGA AAAGGTTC	9527
1253	ACCUUUGU G UCUCCUCU	1614	AGAGGAGA GGCTAGCTACAACGA ACAAAGGT	9528
1294	AGCCGCUU G UUUUGCUC	1615	GAGCAAAA GGCTAGCTACAACGA AAGCGGCT	9529
1344	ACAAUUCU G UCGUGCUC	1616	GAGCACGA GGCTAGCTACAACGA AGAATTGT	9530
1390	GCUAGGCU G UGCUGCCA	1617	TGGCAGCA GGCTAGCTACAACGA AGCCTAGC	9531
1425	CGUCCUUU G UUUACGUC	1618	GACGTAAA GGCTAGCTACAACGA AAAGGACG	9532
1508	CGCCUAUU G UACCGACC	1619	GGTCGGTA GGCTAGCTACAACGA AATAGGCG	9533
1557	CCCCGUCU G UGCCUUCU	1620	AGAAGGCA GGCTAGCTACAACGA AGACGGGG	9534
1581	CGGACCGU G UGCACUUC	1621	GAAGTGCA GGCTAGCTACAACGA ACGGTCCG	9535
1684	UCAGCAAU G UCAACGAC	1622	GTCGTTGA GGCTAGCTACAACGA ATTGCTGA	9536
1719	CAAAGACU G UGUGUUUA	1623	TAAACACA GGCTAGCTACAACGA AGTCTTTG	9537
1721	AAGACUGU G UGUUUAAU	1624	ATTAAACA GGCTAGCTACAACGA ACAGTCTT	9538
1723	GACUGUGU G UUUAAUGA	1625	TCATTAAA GGCTAGCTACAACGA ACACAGTC	9539
1772	AGGUCUUU G UACUAGGA	1626	TCCTAGTA GGCTAGCTACAACGA AAAGACCT	9540
1785	AGGAGGCU G UAGGCAUA	1627	TATGCCTA GGCTAGCTACAACGA AGCCTCCT	9541
1801	AAAUUGGU G UGUUCACC	1628	GGTGAACA GGCTAGCTACAACGA ACCAATTT	9542
1803	AUUGGUGU G UUCACCAG	1629	CTGGTGAA GGCTAGCTACAACGA ACACCAAT	9543
1850	CAUCUCAU G UUCAUGUC	1630	GACATGAA GGCTAGCTACAACGA ATGAGATG	9544

1856	AUGUUCAU G UCCUACUG	1631	CAGTAGGA GGCTAGCTACAACGA ATGAACAT	9545
1864	GUCCUACU G UUCAAGCC	1632	GGCTTGAA GGCTAGCTACAACGA AGTAGGAC	9546
1881	UCCAAGCU G UGCCUUGG	1633	CCAAGGCA GGCTAGCTACAACGA AGCTTGGA	9547
1939	GAGCUUCU G UGGAGUUA	1634	TAACTCCA GGCTAGCTACAACGA AGAAGCTC	9548
2013	UCUGCUCU G UAUCGGGG	1635	CCCCGATA GGCTAGCTACAACGA AGAGCAGA	9549
2045	GGAACAUU G UUCACCUC	1636	GAGGTGAA GGCTAGCTACAACGA AATGTTCC	9550
2082	GCUAUUCU G UGUUGGGG	1637	CCCCAACA GGCTAGCTACAACGA AGAATAGC	9551
2084	UAUUCUGU G UUGGGGUG	1638	CACCCCAA GGCTAGCTACAACGA ACAGAATA	9552
2167	UCAGCUAU G UCAACGUU	1639	AACGTTGA GGCTAGCTACAACGA ATAGCTGA	9553
2205	CAACUAUU G UGGUUUCA	1640	TGAAACCA GGCTAGCTACAACGA AATAGTTG	9554
2222	CAUUUCCU G UCUUACUU	1641	AAGTAAGA GGCTAGCTACAACGA AGGAAATG	9555
2245	GAGAAACU G UUCUUGAA	1642	TTCAAGAA GGCTAGCTACAACGA AGTTTCTC	9556
2262	UAUUUGGU G UCUUUUGG	1643	CCAAAAGA GGCTAGCTACAACGA ACCAAATA	9557
2274	UUUGGAGU G UGGAUUCG	1644	CGAATCCA GGCTAGCTACAACGA ACTCCAAA	9558
2344	AAACUACU G UUGUUAGA	1645	TCTAACAA GGCTAGCTACAACGA AGTAGTTT	9559
2347	CUACUGUU G UUAGACGA	1646	TCGTCTAA GGCTAGCTACAACGA AACAGTAG	9560
2450	AUCUCAAU G UUAGUAUU	1647	AATACTAA GGCTAGCTACAACGA ATTGAGAT	9561
2573	AGGACAUU G UUGAUAGA	1648	TCTATCAA GGCTAGCTACAACGA AATGTCCT	9562
2583	UGAUAGAU G UAAGCAAU	1649	ATTGCTTA GGCTAGCTACAACGA ATCTATCA	9563
2594	AGCAAUUU G UGGGGCCC	1650	GGGCCCCA GGCTAGCTACAACGA AAATTGCT	9564
2663	AUCCCAAU G UUACUAAA	1651	TTTAGTAA GGCTAGCTACAACGA ATTGGGAT	9565
2717	CAGAGUAU G UAGUUAAU	1652	ATTAACTA GGCTAGCTACAACGA ATACTCTG	9566
2901	AUCUUUCU G UCCCCAAU	1653	ATTGGGGA GGCTAGCTACAACGA AGAAAGAT	9567
3071	GGGGGACU G UUGGGGUG	1654	CACCCCAA GGCTAGCTACAACGA AGTCCCCC	9568
3111	UCACAACU G UGCCAGCA	1655	TGCTGGCA GGCTAGCTACAACGA AGTTGTGA	9569
40	AUCCCAGA G UCAGGGCC	1656	GGCCCTGA GGCTAGCTACAACGA TCTGGGAT	9570
46	GAGUCAGG G CCCUGUAC	1657	GTACAGGG GGCTAGCTACAACGA CCTGACTC	9571
65	UCCUGCUG G UGGCUCCA	1658	TGGAGCCA GGCTAGCTACAACGA CAGCAGGA	9572
68	UGCUGGUG G CUCCAGUU	1659	AACTGGAG GGCTAGCTACAACGA CACCAGCA	9573
74	UGGCUCCA G UUCAGGAA	1660	TTCCTGAA GGCTAGCTACAACGA TGGAGCCA	9574
85	CAGGAACA G UGAGCCCU	1661	AGGGCTCA GGCTAGCTACAACGA TGTTCCTG	9575
89	AACAGUGA G CCCUGCUC	1662	GAGCAGGG GGCTAGCTACAACGA TCACTGTT	9576
120	GCCAUAUC G UCAAUCUU	1663	AAGATTGA GGCTAGCTACAACGA GATATGGC	9577
196	CCCUGCUC G UGUUACAG	1664	CTGTAACA GGCTAGCTACAACGA GAGCAGGG	9578
205	UGUUACAG G CGGGGUUU	1665	AAACCCCG GGCTAGCTACAACGA CTGTAACA	9579
210	CAGGCGGG G UUUUUCUU	1666	AAGAAAAA GGCTAGCTACAACGA CCCGCCTG	9580
248	ACCACAGA G UCUAGACU	1667	AGTCTAGA GGCTAGCTACAACGA TCTGTGGT	9581
258	CUAGACUC G UGGUGGAC	1668	GTCCACCA GGCTAGCTACAACGA GAGTCTAG	9582
261	GACUCGUG G UGGACUUC	1669	GAAGTCCA GGCTAGCTACAACGA CACGAGTC	9583
295	GAACACCC G UGUGUCUU	1670	AAGACACA GGCTAGCTACAACGA GGGTGTTC	9584
305	GUGUCUUG G CCAAAAUU	1671	AATTTTGG GGCTAGCTACAACGA CAAGACAC	9585
318	AAUUCGCA G UCCCAAAU	1672	ATTTGGGA GGCTAGCTACAACGA TGCGAATT	9586
332	AAUCUCCA G UCACUCAC	1673	GTGAGTGA GGCTAGCTACAACGA TGGAGATT	9587
368	UUGUCCUG G UUAUCGCU	1674	AGCGATAA GGCTAGCTACAACGA CAGGACAA	9588
390	UGUCUGCG G CGUUUUAU	1675	ATAAAACG GGCTAGCTACAACGA CGCAGACA	9589
392	UCUGCGGC G UUUUAUCA	1676	TGATAAAA GGCTAGCTACAACGA GCCGCAGA	9590
442	UCUUGUUG G UUCUUCUG	1677	CAGAAGAA GGCTAGCTACAACGA CAACAAGA	9591
461	CUAUCAAG G UAUGUUGC	1678	GCAACATA GGCTAGCTACAACGA CTTGATAG	9592
472	UGUUGCCC G UUUGUCCU	1679	AGGACAAA GGCTAGCTACAACGA GGGCAACA	9593
506	AACAACCA G CACCGGAC	1680	GTCCGGTG GGCTAGCTACAACGA TGGTTGTT	9594
625	CAUCUUGG G CUUUCGCA	1681	TGCGAAAG GGCTAGCTACAACGA CCAAGATG	9595

648	CUAUGGGA G UGGGCCUC	1682	GAGGCCCA GGCTAGCTACAACGA TCCCATAG	9596
652	GGGAGUGG G CCUCAGUC	1683	GACTGAGG GGCTAGCTACAACGA CCACTCCC	9597
658	GGGCCUCA G UCCGUUUC	1684	GAAACGGA GGCTAGCTACAACGA TGAGGCCC	9598
662	CUCAGUCC G UUUCUCUU	1685	AAGAGAAA GGCTAGCTACAACGA GGACTGAG	9599
672	UUCUCUUG G CUCAGUUU	1686	AAACTGAG GGCTAGCTACAACGA CAAGAGAA	9600
677	UUGGCUCA G UUUACUAG	1687	CTAGTAAA GGCTAGCTACAACGA TGAGCCAA	9601
685	GUUUACUA G UGCCAUUU	1688	AAATGGCA GGCTAGCTACAACGA TAGTAAAC	9602
699	UUUGUUCA G UGGUUCGU	1689	ACGAACCA GGCTAGCTACAACGA TGAACAAA	9603
702	GUUCAGUG G UUCGUAGG	1690	CCTACGAA GGCTAGCTACAACGA CACTGAAC	9604
706	AGUGGUUC G UAGGGCUU	1691	AAGCCCTA GGCTAGCTACAACGA GAACCACT	9605
711	UUCGUAGG G CUUUCCCC	1692	GGGGAAAG GGCTAGCTACAACGA CCTACGAA	9606
729	ACUGUCUG G CUUUCAGU	1693	ACTGAAAG GGCTAGCTACAACGA CAGACAGT	9607
736	GGCUUUCA G UUAUAUGG	1694	CCATATAA GGCTAGCTACAACGA TGAAAGCC	9608
753	AUGAUGUG G UUUUGGGG	1695	CCCCAAAA GGCTAGCTACAACGA CACATCAT	9609
762	UUUUGGGG G CCAAGUCU	1696	AGACTTGG GGCTAGCTACAACGA CCCCAAAA	9610
767	GGGGCCAA G UCUGUACA	1697	TGTACAGA GGCTAGCTACAACGA TTGGCCCC	9611
785	CAUCUUGA G UCCCUUUA	1698	TAAAGGGA GGCTAGCTACAACGA TCAAGATG	9612
826	GUCUUUGG G UAUACAUU	1699	AATGTATA GGCTAGCTACAACGA CCAAAGAC	9613
898	AAUUGGGA G UUGGGGCA	1700	TGCCCCAA GGCTAGCTACAACGA TCCCAATT	9614
904	GAGUUGGG G CACAUUGC	1701	GCAATGTG GGCTAGCTACAACGA CCCAACTC	9615
971	GUAAACAG G CCUAUUGA	1702	TCAATAGG GGCTAGCTACAACGA CTGTTTAC	9616
987	AUUGGAAA G UAUGUCAA	1703	TTGACATA GGCTAGCTACAACGA TTTCCAAT	9617
1006	AAUUGUGG G UCUUUUGG	1704	CCAAAAGA GGCTAGCTACAACGA CCACAATT	9618
1016	CUUUUGGG G UUUGCCGC	1705	GCGGCAAA GGCTAGCTACAACGA CCCAAAAG	9619
1080	GCAUACAA G CAAAACAG	1706	CTGTTTTG GGCTAGCTACAACGA TTGTATGC	9620
1089	CAAAACAG G CUUUUACU	1707	AGTAAAAG GGCTAGCTACAACGA CTGTTTTG	9621
1116	CUUACAAG G CCUUUCUA	1708	TAGAAAGG GGCTAGCTACAACGA CTTGTAAG	9622
1126	CUUUCUAA G UAAACAGU	1709	ACTGTTTA GGCTAGCTACAACGA TTAGAAAG	9623
1133	AGUAAACA G UAUGUGAA	1710	TTCACATA GGCTAGCTACAACGA TGTTTACT	9624
1152	UUUACCCC G UUGCUCGG	1711	CCGAGCAA GGCTAGCTACAACGA GGGGTAAA	9625
1160	GUUGCUCG G CAACGGCC	1712	GGCCGTTG GGCTAGCTACAACGA CGAGCAAC	9626
1166	CGGCAACG G CCUGGUCU	1713	AGACCAGG GGCTAGCTACAACGA CGTTGCCG	9627
1171	ACGGCCUG G UCUAUGCC	1714	GGCATAGA GGCTAGCTACAACGA CAGGCCGT	9628
1182	UAUGCCAA G UGUUUGCU	1715	AGCAAACA GGCTAGCTACAACGA TTGGCATA	9629
1207	CCCCACUG G UUGGGGCU	1716	AGCCCCAA GGCTAGCTACAACGA CAGTGGGG	9630
1213	UGGUUGGG G CUUGGCCA	1717	TGGCCAAG GGCTAGCTACAACGA CCCAACCA	9631
1218	GGGGCUUG G CCAUAGGC	1718	GCCTATGG GGCTAGCTACAACGA CAAGCCCC	9632
1225	GGCCAUAG G CCAUCAGC	1719	GCTGATGG GGCTAGCTACAACGA CTATGGCC	9633
1232	GGCCAUCA G CGCAUGCG	1720	CGCATGCG GGCTAGCTACAACGA TGATGGCC	9634
1240	GCGCAUGC G UGGAACCU	1721	AGGTTCCA GGCTAGCTACAACGA GCATGCGC	9635
1287	AACUCCUA G CCGCUUGU	1722	ACAAGCGG GGCTAGCTACAACGA TAGGAGTT	9636
1306	UGCUCGCA G CAGGUCUG	1723	CAGACCTG GGCTAGCTACAACGA TGCGAGCA	9637
1310	CGCAGCAG G UCUGGGGC	1724	GCCCCAGA GGCTAGCTACAACGA CTGCTGCG	9638
1317	GGUCUGGG G CAAAACUC	1725	GAGTTTTG GGCTAGCTACAACGA CCCAGACC	9639
1347	AUUCUGUC G UGCUCUCC	1726	GGAGAGCA GGCTAGCTACAACGA GACAGAAT	9640
1379	UUUCCAUG G CUGCUAGG	1727	CCTAGCAG GGCTAGCTACAACGA CATGGAAA	9641
1387	GCUGCUAG G CUGUGCUG	1728	CAGCACAG GGCTAGCTACAACGA CTAGCAGC	9642
1418	CGCGGGAC G UCCUUUGU	1729	ACAAAGGA GGCTAGCTACAACGA GTCCCGCG	9643
1431	UUGUUUAC G UCCCGUCG	1730	CGACGGGA GGCTAGCTACAACGA GTAAACAA	9644
1436	UACGUCCC G UCGGCGCU	1731	AGCGCCGA GGCTAGCTACAACGA GGGACGTA	9645
1440	UCCCGUCG G CGCUGAAU	1732	ATTCAGCG GGCTAGCTACAACGA CGACGGGA	9646

1471	CUCCCGGG G CCGCUUGG	1733	CCAAGCGG GGCTAGCTACAACGA CCCGGGAG	9647
1481	CGCUUGGG G CUCUACCG	1734	CGGTAGAG GGCTAGCTACAACGA CCCAAGCG	9648
1517	UACCGACC G UCCACGGG	1735	CCCGTGGA GGCTAGCTACAACGA GGTCGGTA	9649
1526	UCCACGGG G CGCACCUC	1736	GAGGTGCG GGCTAGCTACAACGA CCCGTGGA	9650
1553	GACUCCCC G UCUGUGCC	1737	GGCACAGA GGCTAGCTACAACGA GGGGAGTC	9651
1579	GCCGGACC G UGUGCACU	1738	AGTGCACA GGCTAGCTACAACGA GGTCCGGC	9652
1605	CUCUGCAC G UCGCAUGG	1739	CCATGCGA GGCTAGCTACAACGA GTGCAGAG	9653
1622	AGACCACC G UGAACGCC	1740	GGCGTTCA GGCTAGCTACAACGA GGTGGTCT	9654
1649	UGCCCAAG G UCUUGCAU	1741	ATGCAAGA GGCTAGCTACAACGA CTTGGGCA	9655
1679	GACUUUCA G CAAUGUCA	1742	TGACATTG GGCTAGCTACAACGA TGAAAGTC	9656
1703	ACCUUGAG G CAUACUUC	1743	GAAGTATG GGCTAGCTACAACGA CTCAAGGT	9657
1732	UUUAAUGA G UGGGAGGA	1744	TCCTCCCA GGCTAGCTACAACGA TCATTAAA	9658
1741	UGGGAGGA G UUGGGGGA	1745	TCCCCCAA GGCTAGCTACAACGA TCCTCCCA	9659
1754	GGGAGGAG G UUAGGUUA	1746	TAACCTAA GGCTAGCTACAACGA CTCCTCCC	9660
1759	GAGGUUAG G UUAAAGGU	1747	ACCTTTAA GGCTAGCTACAACGA CTAACCTC	9661
1766	GGUUAAAG G UCUUUGUA	1748	TACAAAGA GGCTAGCTACAACGA CTTTAACC	9662
1782	ACUAGGAG G CUGUAGGC	1749	GCCTACAG GGCTAGCTACAACGA CTCCTAGT	9663
1789	GGCUGUAG G CAUAAAUU	1750	AATTTATG GGCTAGCTACAACGA CTACAGCC	9664
1799	AUAAAUUG G UGUGUUCA	1751	TGAACACA GGCTAGCTACAACGA CAATTTAT	9665
1811	GUUCACCA G CACCAUGC	1752	GCATGGTG GGCTAGCTACAACGA TGGTGAAC	9666
1870	CUGUUCAA G CCUCCAAG	1753	CTTGGAGG GGCTAGCTACAACGA TTGAACAG	9667
1878	GCCUCCAA G CUGUGCCU	1754	AGGCACAG GGCTAGCTACAACGA TTGGAGGC	9668
1890	UGCCUUGG G UGGCUUUG	1755	CAAAGCCA GGCTAGCTACAACGA CCAAGGCA	9669
1893	CUUGGGUG G CUUUGGGG	1756	CCCCAAAG GGCTAGCTACAACGA CACCCAAG	9670
1901	GCUUUGGG G CAUGGACA	1757	TGTCCATG GGCTAGCTACAACGA CCCAAAGC	9671
1917	AUUGACCC G UAUAAAGA	1758	TCTTTATA GGCTAGCTACAACGA GGGTCAAT	9672
1933	AAUUUGGA G CUUCUGUG	1759	CACAGAAG GGCTAGCTACAACGA TCCAAATT	9673
1944	UCUGUGGA G UUACUCUC	1760	GAGAGTAA GGCTAGCTACAACGA TCCACAGA	9674
2023	AUCGGGGG G CCUUAGAG	1761	CTCTAAGG GGCTAGCTACAACGA CCCCCGAT	9675
2031	GCCUUAGA G UCUCCGGA	1762	TCCGGAGA GGCTAGCTACAACGA TCTAAGGC	9676
2062	ACCAUACG G CACUCAGG	1763	CCTGAGTG GGCTAGCTACAACGA CGTATGGT	9677
2070	GCACUCAG G CAAGCUAU	1764	ATAGCTTG GGCTAGCTACAACGA CTGAGTGC	9678
2074	UCAGGCAA G CUAUUCUG	1765	CAGAATAG GGCTAGCTACAACGA TTGCCTGA	9679
2090	GUGUUGGG G UGAGUUGA	1766	TCAACTCA GGCTAGCTACAACGA CCCAACAC	9680
2094	UGGGGUGA G UUGAUGAA	1767	TTCATCAA GGCTAGCTACAACGA TCACCCCA	9681
2107	UGAAUCUA G CCACCUGG	1768	CCAGGTGG GGCTAGCTACAACGA TAGATTCA	9682
2116	CCACCUGG G UGGGAAGU	1769	ACTTCCCA GGCTAGCTACAACGA CCAGGTGG	9683
2123	GGUGGGAA G UAAUUUGG	1770	CCAAATTA GGCTAGCTACAACGA TTCCCACC	9684
2140	AAGAUCCA G CAUCCAGG	1771	CCTGGATG GGCTAGCTACAACGA TGGATCTT	9685
2155	GGGAAUUA G UAGUCAGC	1772	GCTGACTA GGCTAGCTACAACGA TAATTCCC	9686
2158	AAUUAGUA G UCAGCUAU	1773	ATAGCTGA GGCTAGCTACAACGA TACTAATT	9687
2162	AGUAGUCA G CUAUGUCA	1774	TGACATAG GGCTAGCTACAACGA TGACTACT	9688
2173	AUGUCAAC G UUAAUAUG	1775	CATATTAA GGCTAGCTACAACGA GTTGACAT	9689
2183	UAAUAUGG G CCUAAAAA	1776	TTTTTAGG GGCTAGCTACAACGA CCATATTA	9690
2208	CUAUUGUG G UUUCACAU	1777	ATGTGAAA GGCTAGCTACAACGA CACAATAG	9691
2235	ACUUUUGG G CGAGAAAC	1778	GTTTCTCG GGCTAGCTACAACGA CCAAAAGT	9692
2260	AAUAUUUG G UGUCUUUU	1779	AAAAGACA GGCTAGCTACAACGA CAAATATT	9693
2272	CUUUUGGA G UGUGGAUU	1780	AATCCACA GGCTAGCTACAACGA TCCAAAAG	9694
2360	ACGAAGAG G CAGGUCCC	1781	GGGACCTG GGCTAGCTACAACGA CTCTTCGT	9695
2364	AGAGGCAG G UCCCCUAG	1782	CTAGGGGA GGCTAGCTACAACGA CTGCCTCT	9696
2403	AGACGAAG G UCUCAAUC	1783	GATTGAGA GGCTAGCTACAACGA CTTCGTCT	9697

2417	AUCGCCGC G UCGCAGAA	1784	TTCTGCGA GGCTAGCTACAACGA GCGGCGAT	9698
2454	CAAUGUUA G UAUUCCUU	1785	AAGGAATA GGCTAGCTACAACGA TAACATTG	9699
2474	CACAUAAG G UGGGAAAC	1786	GTTTCCCA GGCTAGCTACAACGA CTTATGTG	9700
2491	UUUACGGG G CUUUAUUC	1787	GAATAAAG GGCTAGCTACAACGA CCCGTAAA	9701
2507	CUUCUACG G UACCUUGC	1788	GCAAGGTA GGCTAGCTACAACGA CGTAGAAG	9702
2530	CCUAAAUG G CAAACUCC	1789	GGAGTTTG GGCTAGCTACAACGA CATTTAGG	9703
2587	AGAUGUAA G CAAUUUGU	1790	ACAAATTG GGCTAGCTACAACGA TTACATCT	9704
2599	UUUGUGGG G CCCCUUAC	1791	GTAAGGGG GGCTAGCTACAACGA CCCACAAA	9705
2609	CCCUUACA G UAAAUGAA	1792	TTCATTTA GGCTAGCTACAACGA TGTAAGGG	9706
2650	CCUGCUAG G UUUUAUCC	1793	GGATAAAA GGCTAGCTACAACGA CTAGCAGG	9707
2701	AUCAAACC G UAUUAUCC	1794	GGATAATA GGCTAGCTACAACGA GGTTTGAT	9708
2713	UAUCCAGA G UAUGUAGU	1795	ACTACATA GGCTAGCTACAACGA TCTGGATA	9709
2720	AGUAUGUA G UUAAUCAU	1796	ATGATTAA GGCTAGCTACAACGA TACATACT	9710
2768	UUUGGAAG G CGGGGAUC	1797	GATCCCCG GGCTAGCTACAACGA CTTCCAAA	9711
2791	AAAAGAGA G UCCACACG	1798	CGTGTGGA GGCTAGCTACAACGA TCTCTTTT	9712
2799	GUCCACAC G UAGCGCCU	1799	AGGCGCTA GGCTAGCTACAACGA GTGTGGAC	9713
2802	CACACGUA G CGCCUCAU	1800	ATGAGGCG GGCTAGCTACAACGA TACGTGTG	9714
2818	UUUUGCGG G UCACCAUA	1801	TATGGTGA GGCTAGCTACAACGA CCGCAAAA	9715
2848	GAUCUACA G CAUGGGAG	1802	CTCCCATG GGCTAGCTACAACGA TGTAGATC	9716
2857	CAUGGGAG G UUGGUCUU	1803	AAGACCAA GGCTAGCTACAACGA CTCCCATG	9717
2861	GGAGGUUG G UCUUCCAA	1804	TTGGAAGA GGCTAGCTACAACGA CAACCTCC	9718
2881	UCGAAAAG G CAUGGGGA	1805	TCCCCATG GGCTAGCTACAACGA CTTTTCGA	9719
2936	GAUCAUCA G UUGGACCC	1806	GGGTCCAA GGCTAGCTACAACGA TGATGATC	9720
2955	CAUUCAAA G CCAACUCA	1807	TGAGTTGG GGCTAGCTACAACGA TTTGAATG	9721
2964	CCAACUCA G UAAAUCCA	1808	TGGATTTA GGCTAGCTACAACGA TGAGTTGG	9722
3005	GACAACUG G CCGGACGC	1809	GCGTCCGG GGCTAGCTACAACGA CAGTTGTC	9723
3021	CCAACAAG G UGGGAGUG	1810	CACTCCCA GGCTAGCTACAACGA CTTGTTGG	9724
3027	AGGUGGGA G UGGGAGCA	1811	TGCTCCCA GGCTAGCTACAACGA TCCCACCT	9725
3033	GAGUGGGA G CAUUCGGG	1812	CCCGAATG GGCTAGCTACAACGA TCCCACTC	9726
3041	GCAUUCGG G CCAGGGUU	1813	AACCCTGG GGCTAGCTACAACGA CCGAATGC	9727
3047	GGGCCAGG G UUCACCCC	1814	GGGGTGAA GGCTAGCTACAACGA CCTGGCCC	9728
3077	CUGUUGGG G UGGAGCCC	1815	GGGCTCCA GGCTAGCTACAACGA CCCAACAG	9729
3082	GGGGUGGA G CCCUCACG	1816	CGTGAGGG GGCTAGCTACAACGA TCCACCCC	9730
3097	CGCUCAGG G CCUACUCA	1817	TGAGTAGG GGCTAGCTACAACGA CCTGAGCG	9731
3117	CUGUGCCA G CAGCUCCU	1818	AGGAGCTG GGCTAGCTACAACGA TGGCACAG	9732
3120	UGCCAGCA G CUCCUCCU	1819	AGGAGGAG GGCTAGCTACAACGA TGCTGGCA	9733
3146	ACCAAUCG G CAGUCAGG	1820	CCTGACTG GGCTAGCTACAACGA CGATTGGT	9734
3149	AAUCGGCA G UCAGGAAG	1821	CTTCCTGA GGCTAGCTACAACGA TGCCGATT	9735
3158	UCAGGAAG G CAGCCUAC	1822	GTAGGCTG GGCTAGCTACAACGA CTTCCTGA	9736
3161	GGAAGGCA G CCUACUCC	1823	GGAGTAGG GGCTAGCTACAACGA TGCCTTCC	9737
3204	AUCCUCAG G CCAUGCAG	1824	CTGCATGG GGCTAGCTACAACGA CTGAGGAT	9738
10	ACUCCACC A CUUUCCAC	703	GTGGAAAG GGCTAGCTACAACGA GGTGGAGT	9739
17	CACUUUCC A CCAAACUC	706	GAGTTTGG GGCTAGCTACAACGA GGAAAGTG	9740
22	UCCACCAA A CUCUUCAA	1825	TTGAAGAG GGCTAGCTACAACGA TTGGTGGA	9741
32	UCUUCAAG A UCCCAGAG	1826	CTCTGGGA GGCTAGCTACAACGA CTTGAAGA	9742
53	GGCCCUGU A CUUUCCUG	42	CAGGAAAG GGCTAGCTACAACGA ACAGGGCC	9743
82	GUUCAGGA A CAGUGAGC	1827	GCTCACTG GGCTAGCTACAACGA TCCTGAAC	9744
101	UGCUCAGA A UACUGUCU	1828	AGACAGTA GGCTAGCTACAACGA TCTGAGCA	9745
103	CUCAGAAU A CUGUCUCU	50	AGAGACAG GGCTAGCTACAACGA ATTCTGAG	9746
115	UCUCUGCC A UAUCGUCA	737	TGACGATA GGCTAGCTACAACGA GGCAGAGA	9747
117	UCUGCCAU A UCGUCAAU	53	ATTGACGA GGCTAGCTACAACGA ATGGCAGA	9748

124	UAUCGUCA A UCUUAUCG	1829	CGATAAGA GGCTAGCTACAACGA TGACGATA	9749
129	UCAAUCUU A UCGAAGAC	58	GTCTTCGA GGCTAGCTACAACGA AAGATTGA	9750
136	UAUCGAAG A CUGGGGAC	1830	GTCCCCAG GGCTAGCTACAACGA CTTCGATA	9751
143	GACUGGGG A CCCUGUAC	1831	GTACAGGG GGCTAGCTACAACGA CCCCAGTC	9752
150	GACCCUGU A CCGAACAU	60	ATGTTCGG GGCTAGCTACAACGA ACAGGGTC	9753
155	UGUACCGA A CAUGGAGA	1832	TCTCCATG GGCTAGCTACAACGA TCGGTACA	9754
157	UACCGAAC A UGGAGAAC	745	GTTCTCCA GGCTAGCTACAACGA GTTCGGTA	9755
164	CAUGGAGA A CAUCGCAU	1833	ATGCGATG GGCTAGCTACAACGA TCTCCATG	9756
166	UGGAGAAC A UCGCAUCA	746	TGATGCGA GGCTAGCTACAACGA GTTCTCCA	9757
171	AACAUCGC A UCAGGACU	747	AGTCCTGA GGCTAGCTACAACGA GCGATGTT	9758
177	GCAUCAGG A CUCCUAGG	1834	CCTAGGAG GGCTAGCTACAACGA CCTGATGC	9759
186	CUCCUAGG A CCCCUGCU	1835	AGCAGGGG GGCTAGCTACAACGA CCTAGGAG	9760
201	CUCGUGUU A CAGGCGGG	67	CCCGCCTG GGCTAGCTACAACGA AACACGAG	9761
223	UCUUGUUG A CAAAAAUC	1836	GATTTTTG GGCTAGCTACAACGA CAACAAGA	9762
229	UGACAAAA A UCCUCACA	1837	TGTGAGGA GGCTAGCTACAACGA TTTTGTCA	9763
235	AAAUCCUC A CAAUACCA	762	TGGTATTG GGCTAGCTACAACGA GAGGATTT	9764
238	UCCUCACA A UACCACAG	1838	CTGTGGTA GGCTAGCTACAACGA TGTGAGGA	9765
240	CUCACAAU A CCACAGAG	77	CTCTGTGG GGCTAGCTACAACGA ATTGTGAG	9766
243	ACAAUACC A CAGAGUCU	765	AGACTCTG GGCTAGCTACAACGA GGTATTGT	9767
254	GAGUCUAG A CUCGUGGU	1839	ACCACGAG GGCTAGCTACAACGA CTAGACTC	9768
265	CGUGGUGG A CUUCUCUC	1840	GAGAGAAG GGCTAGCTACAACGA CCACCACG	9769
275	UUCUCUCA A UUUUCUAG	1841	CTAGAAAA GGCTAGCTACAACGA TGAGAGAA	9770
289	UAGGGGGA A CACCCGUG	1842	CACGGGTG GGCTAGCTACAACGA TCCCCCTA	9771
291	GGGGGAAC A CCCGUGUG	774	CACACGGG GGCTAGCTACAACGA GTTCCCCC	9772
311	UGGCCAAA A UUCGCAGU	1843	ACTGCGAA GGCTAGCTACAACGA TTTGGCCA	9773
325	AGUCCCAA A UCUCCAGU	1844	ACTGGAGA GGCTAGCTACAACGA TTGGGACT	9774
335	CUCCAGUC A CUCACCAA	787	TTGGTGAG GGCTAGCTACAACGA GACTGGAG	9775
339	AGUCACUC A CCAACCUG	789	CAGGTTGG GGCTAGCTACAACGA GAGTGACT	9776
343	ACUCACCA A CCUGUUGU	1845	ACAACAGG GGCTAGCTACAACGA TGGTGAGT	9777
358	GUCCUCCA A UUUGUCCU	1846	AGGACAAA GGCTAGCTACAACGA TGGAGGAC	9778
371	UCCUGGUU A UCGCUGGA	106	TCCAGCGA GGCTAGCTACAACGA AACCAGGA	9779
379	AUCGCUGG A UGUGUCUG	1847	CAGACACA GGCTAGCTACAACGA CCAGCGAT	9780
397	GGCGUUUU A UCAUCUUC	112	GAAGATGA GGCTAGCTACAACGA AAAACGCC	9781
400	GUUUUAUC A UCUUCCUC	802	GAGGAAGA GGCTAGCTACAACGA GATAAAAC	9782
412	UCCUCUGC A UCCUGCUG	807	CAGCAGGA GGCTAGCTACAACGA GCAGAGGA	9783
423	CUGCUGCU A UGCCUCAU	119	ATGAGGCA GGCTAGCTACAACGA AGCAGCAG	9784
430	UAUGCCUC A UCUUCUUG	814	CAAGAAGA GGCTAGCTACAACGA GAGGCATA	9785
452	UCUUCUGG A CUAUCAAG	1848	CTTGATAG GGCTAGCTACAACGA CCAGAAGA	9786
455	UCUGGACU A UCAAGGUA	130	TACCTTGA GGCTAGCTACAACGA AGTCCAGA	9787
463	AUCAAGGU A UGUUGCCC	132	GGGCAACA GGCTAGCTACAACGA ACCTTGAT	9788
484	GUCCUCUA A UUCCAGGA	1849	TCCTGGAA GGCTAGCTACAACGA TAGAGGAC	9789
492	AUUCCAGG A UCAUCAAC	1850	GTTGATGA GGCTAGCTACAACGA CCTGGAAT	9790
495	CCAGGAUC A UCAACAAC	828	GTTGTTGA GGCTAGCTACAACGA GATCCTGG	9791
499	GAUCAUCA A CAACCAGC	1851	GCTGGTTG GGCTAGCTACAACGA TGATGATC	9792
502	CAUCAACA A CCAGCACC	1852	GGTGCTGG GGCTAGCTACAACGA TGTTGATG	9793
513	AGCACCGG A CCAUGCAA	1853	TTGCATGG GGCTAGCTACAACGA CCGGTGCT	9794
516	ACCGGACC A UGCAAAAC	836	GTTTTGCA GGCTAGCTACAACGA GGTCCGGT	9795
523	CAUGCAAA A CCUGCACA	1854	TGTGCAGG GGCTAGCTACAACGA TTTGCATG	9796
529	AAACCUGC A CAACUCCU	840	AGGAGTTG GGCTAGCTACAACGA GCAGGTTT	9797
532	CCUGCACA A CUCCUGCU	1855	AGCAGGAG GGCTAGCTACAACGA TGTGCAGG	9798
547	CUCAAGGA A CCUCUAUG	1856	CATAGAGG GGCTAGCTACAACGA TCCTTGAG	9799

553	GAACCUCU A UGUUUCCC	146	GGGAAACA GGCTAGCTACAACGA AGAGGTTC	9800
564	UUUCCCUC A UGUUGCUG	853	CAGCAACA GGCTAGCTACAACGA GAGGGAAA	9801
574	GUUGCUGU A CAAAACCU	152	AGGTTTTG GGCTAGCTACAACGA ACAGCAAC	9802
579	UGUACAAA A CCUACGGA	1857	TCCGTAGG GGCTAGCTACAACGA TTTGTACA	9803
583	CAAAACCU A CGGACGGA	153	TCCGTCCG GGCTAGCTACAACGA AGGTTTTG	9804
587	ACCUACGG A CGGAAACU	1858	AGTTTCCG GGCTAGCTACAACGA CCGTAGGT	9805
593	GGACGGAA A CUGCACCU	1859	AGGTGCAG GGCTAGCTACAACGA TTCCGTCC	9806
598	GAAACUGC A CCUGUAUU	859	AATACAGG GGCTAGCTACAACGA GCAGTTTC	9807
604	GCACCUGU A UUCCCAUC	154	GATGGGAA GGCTAGCTACAACGA ACAGGTGC	9808
610	GUAUUCCC A UCCCAUCA	864	TGATGGGA GGCTAGCTACAACGA GGGAATAC	9809
615	CCCAUCCC A UCAUCUUG	867	CAAGATGA GGCTAGCTACAACGA GGGATGGG	9810
618	AUCCCAUC A UCUUGGGC	868	GCCCAAGA GGCTAGCTACAACGA GATGGGAT	9811
636	UUCGCAAA A UACCUAUG	1860	CATAGGTA GGCTAGCTACAACGA TTTGCGAA	9812
638	CGCAAAAU A CCUAUGGG	164	CCCATAGG GGCTAGCTACAACGA ATTTTGCG	9813
642	AAAUACCU A UGGGAGUG	165	CACTCCCA GGCTAGCTACAACGA AGGTATTT	9814
681	CUCAGUUU A CUAGUGCC	176	GGCACTAG GGCTAGCTACAACGA AAACTGAG	9815
690	CUAGUGCC A UUUGUUCA	884	TGAACAAA GGCTAGCTACAACGA GGCACTAG	9816
721	UUUCCCCC A CUGUCUGG	891	CCAGACAG GGCTAGCTACAACGA GGGGGAAA	9817
739	UUUCAGUU A UAUGGAUG	193	CATCCATA GGCTAGCTACAACGA AACTGAAA	9818
741	UCAGUUAU A UGGAUGAU	194	ATCATCCA GGCTAGCTACAACGA ATAACTGA	9819
745	UUAUAUGG A UGAUGUGG	1861	CCACATCA GGCTAGCTACAACGA CCATATAA	9820
748	UAUGGAUG A UGUGGUUU	1862	AAACCACA GGCTAGCTACAACGA CATCCATA	9821
773	AAGUCUGU A CAACAUCU	199	AGATGTTG GGCTAGCTACAACGA ACAGACTT	9822
776	UCUGUACA A CAUCUUGA	1863	TCAAGATG GGCTAGCTACAACGA TGTACAGA	9823
778	UGUACAAC A UCUUGAGU	900	ACTCAAGA GGCTAGCTACAACGA GTTGTACA	9824
793	GUCCCUUU A UGCCGCUG	205	CAGCGGCA GGCTAGCTACAACGA AAAGGGAC	9825
804	CCGCUGUU A CCAAUUUU	207	AAAATTGG GGCTAGCTACAACGA AACAGCGG	9826
808	UGUUACCA A UUUUCUUU	1864	AAAGAAAA GGCTAGCTACAACGA TGGTAACA	9827
828	CUUUGGGU A UACAUUUA	218	TAAATGTA GGCTAGCTACAACGA ACCCAAAG	9828
830	UUGGGUAU A CAUUUAAA	219	TTTAAATG GGCTAGCTACAACGA ATACCCAA	9829
832	GGGUAUAC A UUUAAACC	911	GGTTTAAA GGCTAGCTACAACGA GTATACCC	9830
838	ACAUUUAA A CCCUCACA	1865	TGTGAGGG GGCTAGCTACAACGA TTAAATGT	9831
844	AAACCCUC A CAAAACAA	915	TTGTTTTG GGCTAGCTACAACGA GAGGGTTT	9832
849	CUCACAAA A CAAAAAGA	1866	TCTTTTTG GGCTAGCTACAACGA TTTGTGAG	9833
857	ACAAAAAG A UGGGGAUA	1867	TATCCCCA GGCTAGCTACAACGA CTTTTTGT	9834
863	AGAUGGGG A UAUUCCCU	1868	AGGGAATA GGCTAGCTACAACGA CCCCATCT	9835
865	AUGGGGAU A UUCCCUUA	224	TAAGGGAA GGCTAGCTACAACGA ATCCCCAT	9836
874	UUCCCUUA A CUUCAUGG	1869	CCATGAAG GGCTAGCTACAACGA TAAGGGAA	9837
879	UUAACUUC A UGGGAUAU	922	ATATCCCA GGCTAGCTACAACGA GAAGTTAA	9838
884	UUCAUGGG A UAUGUAAU	1870	ATTACATA GGCTAGCTACAACGA CCCATGAA	9839
886	CAUGGGAU A UGUAAUUG	231	CAATTACA GGCTAGCTACAACGA ATCCCATG	9840
891	GAUAUGUA A UUGGGAGU	1871	ACTCCCAA GGCTAGCTACAACGA TACATATC	9841
906	GUUGGGGC A CAUUGCCA	923	TGGCAATG GGCTAGCTACAACGA GCCCCAAC	9842
908	UGGGGCAC A UUGCCACA	924	TGTGGCAA GGCTAGCTACAACGA GTGCCCCA	9843
914	ACAUUGCC A CAGGAACA	926	TGTTCCTG GGCTAGCTACAACGA GGCAATGT	9844
920	CCACAGGA A CAUAUUGU	1872	ACAATATG GGCTAGCTACAACGA TCCTGTGG	9845
922	ACAGGAAC A UAUUGUAC	928	GTACAATA GGCTAGCTACAACGA GTTCCTGT	9846
924	AGGAACAU A UUGUACAA	236	TTGTACAA GGCTAGCTACAACGA ATGTTCCT	9847
929	CAUAUUGU A CAAAAAAU	238	ATTTTTTG GGCTAGCTACAACGA ACAATATG	9848
936	UACAAAAA A UCAAAAUG	1873	CATTTTGA GGCTAGCTACAACGA TTTTTGTA	9849
942	AAAUCAAA A UGUGUUUU	1874	AAAACACA GGCTAGCTACAACGA TTTGATTT	9850

956	UUUAGGAA A CUUCCUGU	1875	ACAGGAAG GGCTAGCTACAACGA TTCCTAAA	9851
967	UCCUGUAA A CAGGCCUA	1876	TAGGCCTG GGCTAGCTACAACGA TTACAGGA	9852
975	ACAGGCCU A UUGAUUGG	247	CCAATCAA GGCTAGCTACAACGA AGGCCTGT	9853
979	GCCUAUUG A UUGGAAAG	1877	CTTTCCAA GGCTAGCTACAACGA CAATAGGC	9854
989	UGGAAAGU A UGUCAACG	250	CGTTGACA GGCTAGCTACAACGA ACTTTCCA	9855
995	GUAUGUCA A CGAAUUGU	1878	ACAATTCG GGCTAGCTACAACGA TGACATAC	9856
999	GUCAACGA A UUGUGGGU	1879	ACCCACAA GGCTAGCTACAACGA TCGTTGAC	9857
1032	CCCCUUUC A CGCAAUGU	944	ACATTGCG GGCTAGCTACAACGA GAAAGGGG	9858
1037	UUCACGCA A UGUGGAUA	1880	TATCCACA GGCTAGCTACAACGA TGCGTGAA	9859
1043	CAAUGUGG A UAUUCUGC	1881	GCAGAATA GGCTAGCTACAACGA CCACATTG	9860
1045	AUGUGGAU A UUCUGCUU	262	AAGCAGAA GGCTAGCTACAACGA ATCCACAT	9861
1056	CUGCUUUA A UGCCUUUA	1882	TAAAGGCA GGCTAGCTACAACGA TAAAGCAG	9862
1064	AUGCCUUU A UAUGCAUG	270	CATGCATA GGCTAGCTACAACGA AAAGGCAT	9863
1066	GCCUUUAU A UGCAUGCA	271	TGCATGCA GGCTAGCTACAACGA ATAAAGGC	9864
1070	UUAUAUGC A UGCAUACA	950	TGTATGCA GGCTAGCTACAACGA GCATATAA	9865
1074	AUGCAUGC A UACAAGCA	951	TGCTTGTA GGCTAGCTACAACGA GCATGCAT	9866
1076	GCAUGCAU A CAAGCAAA	272	TTTGCTTG GGCTAGCTACAACGA ATGCATGC	9867
1085	CAAGCAAA A CAGGCUUU	1883	AAAGCCTG GGCTAGCTACAACGA TTTGCTTG	9868
1095	AGGCUUUU A CUUUCUCG	276	CGAGAAAG GGCTAGCTACAACGA AAAAGCCT	9869
1107	UCUCGCCA A CUUACAAG	1884	CTTGTAAG GGCTAGCTACAACGA TGGCGAGA	9870
1111	GCCAACUU A CAAGGCCU	282	AGGCCTTG GGCTAGCTACAACGA AAGTTGGC	9871
1130	CUAAGUAA A CAGUAUGU	1885	ACATACTG GGCTAGCTACAACGA TTACTTAG	9872
1135	UAAACAGU A UGUGAACC	288	GGTTCACA GGCTAGCTACAACGA ACTGTTTA	9873
1141	GUAUGUGA A CCUUUACC	1886	GGTAAAGG GGCTAGCTACAACGA TCACATAC	9874
1147	GAACCUUU A CCCCGUUG	291	CAACGGGG GGCTAGCTACAACGA AAAGGTTC	9875
1163	GCUCGGCA A CGGCCUGG	1887	CCAGGCCG GGCTAGCTACAACGA TGCCGAGC	9876
1175	CCUGGUCU A UGCCAAGU	295	ACTTGGCA GGCTAGCTACAACGA AGACCAGG	9877
1192	GUUUGCUG A CGCAACCC	1888	GGGTTGCG GGCTAGCTACAACGA CAGCAAAC	9878
1197	CUGACGCA A CCCCCACU	1889	AGTGGGGG GGCTAGCTACAACGA TGCGTCAG	9879
1203	CAACCCCC A CUGGUUGG	984	CCAACCAG GGCTAGCTACAACGA GGGGGTTG	9880
1221	GCUUGGCC A UAGGCCAU	988	ATGGCCTA GGCTAGCTACAACGA GGCCAAGC	9881
1228	CAUAGGCC A UCAGCGCA	990	TGCGCTGA GGCTAGCTACAACGA GGCCTATG	9882
1236	AUCAGCGC A UGCGUGGA	992	TCCACGCA GGCTAGCTACAACGA GCGCTGAT	9883
1245	UGCGUGGA A CCUUUGUG	1890	CACAAAGG GGCTAGCTACAACGA TCCACGCA	9884
1266	CUCUGCCG A UCCAUACC	1891	GGTATGGA GGCTAGCTACAACGA CGGCAGAG	9885
1270	GCCGAUCC A UACCGCGG	1001	CCGCGGTA GGCTAGCTACAACGA GGATCGGC	9886
1272	CGAUCCAU A CCGCGGAA	308	TTCCGCGG GGCTAGCTACAACGA ATGGATCG	9887
1280	ACCGCGGA A CUCCUAGC	1892	GCTAGGAG GGCTAGCTACAACGA TCCGCGGT	9888
1322	GGGGCAAA A CUCAUCGG	1893	CCGATGAG GGCTAGCTACAACGA TTTGCCCC	9889
1326	CAAAACUC A UCGGGACU	1014	AGTCCCGA GGCTAGCTACAACGA GAGTTTTG	9890
1332	UCAUCGGG A CUGACAAU	1894	ATTGTCAG GGCTAGCTACAACGA CCCGATGA	9891
1336	CGGGACUG A CAAUUCUG	1895	CAGAATTG GGCTAGCTACAACGA CAGTCCCG	9892
1339	GACUGACA A UUCUGUCG	1896	CGACAGAA GGCTAGCTACAACGA TGTCAGTC	9893
1361	UCCCGCAA A UAUACAUC	1897	GATGTATA GGCTAGCTACAACGA TTGCGGGA	9894
1363	CCGCAAAU A UACAUCAU	324	ATGATGTA GGCTAGCTACAACGA ATTTGCGG	9895
1365	GCAAAUAU A CAUCAUUU	325	AAATGATG GGCTAGCTACAACGA ATATTTGC	9896
1367	AAAUAUAC A UCAUUUCC	1023	GGAAATGA GGCTAGCTACAACGA GTATATTT	9897
1370	UAUACAUC A UUUCCAUG	1024	CATGGAAA GGCTAGCTACAACGA GATGTATA	9898
1376	UCAUUUCC A UGGCUGCU	1026	AGCAGCCA GGCTAGCTACAACGA GGAAATGA	9899
1399	UGCUGCCA A CUGGAUCC	1898	GGATCCAG GGCTAGCTACAACGA TGGCAGCA	9900
1404	CCAACUGG A UCCUACGC	1899	GCGTAGGA GGCTAGCTACAACGA CCAGTTGG	9901

1409	UGGAUCCU A CGCGGGAC	332	GTCCCGCG GGCTAGCTACAACGA AGGATCCA	9902
1416	UACGCGGG A CGUCCUUU	1900	AAAGGACG GGCTAGCTACAACGA CCCGCGTA	9903
1429	CUUUGUUU A CGUCCCGU	338	ACGGGACG GGCTAGCTACAACGA AAACAAAG	9904
1447	GGCGCUGA A UCCCGCGG	1901	CCGCGGGA GGCTAGCTACAACGA TCAGCGCC	9905
1456	UCCCGCGG A CGACCCCU	1902	AGGGGTCG GGCTAGCTACAACGA CCGCGGGA	9906
1459	CGCGGACG A CCCCUCCC	1903	GGGAGGGG GGCTAGCTACAACGA CGTCCGCG	9907
1486	GGGGCUCU A CCGCCCGC	345	GCGGGCGG GGCTAGCTACAACGA AGAGCCCC	9908
1505	CUCCGCCU A UUGUACCG	349	CGGTACAA GGCTAGCTACAACGA AGGCGGAG	9909
1510	CCUAUUGU A CCGACCGU	351	ACGGTCGG GGCTAGCTACAACGA ACAATAGG	9910
1514	UUGUACCG A CCGUCCAC	1904	GTGGACGG GGCTAGCTACAACGA CGGTACAA	9911
1521	GACCGUCC A CGGGGCGC	1064	GCGCCCCG GGCTAGCTACAACGA GGACGGTC	9912
1530	CGGGGCGC A CCUCUCUU	1065	AAGAGAGG GGCTAGCTACAACGA GCGCCCCG	9913
1540	CUCUCUUU A CGCGGACU	357	AGTCCGCG GGCTAGCTACAACGA AAAGAGAG	9914
1546	UUACGCGG A CUCCCCGU	1905	ACGGGGAG GGCTAGCTACAACGA CCGCGTAA	9915
1567	GCCUUCUC A UCUGCCGG	1078	CCGGCAGA GGCTAGCTACAACGA GAGAAGGC	9916
1576	UCUGCCGG A CCGUGUGC	1906	GCACACGG GGCTAGCTACAACGA CCGGCAGA	9917
1585	CCGUGUGC A CUUCGCUU	1082	AAGCGAAG GGCTAGCTACAACGA GCACACGG	9918
1595	UUCGCUUC A CCUCUGCA	1085	TGCAGAGG GGCTAGCTACAACGA GAAGCGAA	9919
1603	ACCUCUGC A CGUCGCAU	1089	ATGCGACG GGCTAGCTACAACGA GCAGAGGT	9920
1610	CACGUCGC A UGGAGACC	1090	GGTCTCCA GGCTAGCTACAACGA GCGACGTG	9921
1616	GCAUGGAG A CCACCGUG	1907	CACGGTGG GGCTAGCTACAACGA CTCCATGC	9922
1619	UGGAGACC A CCGUGAAC	1092	GTTCACGG GGCTAGCTACAACGA GGTCTCCA	9923
1626	CACCGUGA A CGCCCACA	1908	TGTGGGCG GGCTAGCTACAACGA TCACGGTG	9924
1638	CCACAGGA A CCUGCCCA	1909	TGGGCAGG GGCTAGCTACAACGA TCCTGTGG	9925
1656	GGUCUUGC A UAAGAGGA	1104	TCCTCTTA GGCTAGCTACAACGA GCAAGACC	9926
1664	AUAAGAGG A CUCUUGGA	1910	TCCAAGAG GGCTAGCTACAACGA CCTCTTAT	9927
1672	ACUCUUGG A CUUUCAGC	1911	GCTGAAAG GGCTAGCTACAACGA CCAAGAGT	9928
1682	UUUCAGCA A UGUCAACG	1912	CGTTGACA GGCTAGCTACAACGA TGCTGAAA	9929
1688	CAAUGUCA A CGACCGAC	1913	GTCGGTCG GGCTAGCTACAACGA TGACATTG	9930
1691	UGUCAACG A CCGACCUU	1914	AAGGTCGG GGCTAGCTACAACGA CGTTGACA	9931
1695	AACGACCG A CCUUGAGG	1915	CCTCAAGG GGCTAGCTACAACGA CGGTCGTT	9932
1705	CUUGAGGC A UACUUCAA	1114	TTGAAGTA GGCTAGCTACAACGA GCCTCAAG	9933
1707	UGAGGCAU A CUUCAAAG	380	CTTTGAAG GGCTAGCTACAACGA ATGCCTCA	9934
1716	CUUCAAAG A CUGUGUGU	1916	ACACACAG GGCTAGCTACAACGA CTTTGAAG	9935
1728	UGUGUUUA A UGAGUGGG	1917	CCCACTCA GGCTAGCTACAACGA TAAACACA	9936
1774	GUCUUUGU A CUAGGAGG	394	CCTCCTAG GGCTAGCTACAACGA ACAAAGAC	9937
1791	CUGUAGGC A UAAAUUGG	1121	CCAATTTA GGCTAGCTACAACGA GCCTACAG	9938
1795	AGGCAUAA A UUGGUGUG	1918	CACACCAA GGCTAGCTACAACGA TTATGCCT	9939
1807	GUGUGUUC A CCAGCACC	1122	GGTGCTGG GGCTAGCTACAACGA GAACACAC	9940
1813	UCACCAGC A CCAUGCAA	1125	TTGCATGG GGCTAGCTACAACGA GCTGGTGA	9941
1816	CCAGCACC A UGCAACUU	1127	AAGTTGCA GGCTAGCTACAACGA GGTGCTGG	9942
1821	ACCAUGCA A CUUUUUCA	1919	TGAAAAAG GGCTAGCTACAACGA TGCATGGT	9943
1829	ACUUUUUC A CCUCUGCC	1130	GGCAGAGG GGCTAGCTACAACGA GAAAAAGT	9944
1840	UCUGCCUA A UCAUCUCA	1920	TGAGATGA GGCTAGCTACAACGA TAGGCAGA	9945
1843	GCCUAAUC A UCUCAUGU	1136	ACATGAGA GGCTAGCTACAACGA GATTAGGC	9946
1848	AUCAUCUC A UGUUCAUG	1138	CATGAACA GGCTAGCTACAACGA GAGATGAT	9947
1854	UCAUGUUC A UGUCCUAC	1139	GTAGGACA GGCTAGCTACAACGA GAACATGA	9948
1861	CAUGUCCU A CUGUUCAA	414	TTGAACAG GGCTAGCTACAACGA AGGACATG	9949
1903	UUUGGGGC A UGGACAUU	1152	AATGTCCA GGCTAGCTACAACGA GCCCCAAA	9950
1907	GGGCAUGG A CAUUGACC	1921	GGTCAATG GGCTAGCTACAACGA CCATGCCC	9951
1909	GCAUGGAC A UUGACCCG	1153	CGGGTCAA GGCTAGCTACAACGA GTCCATGC	9952

1913	GGACAUUG A CCCGUAUA	1922	TATACGGG GGCTAGCTACAACGA CAATGTCC	9953
1919	UGACCCGU A UAAAGAAU	422	ATTCTTTA GGCTAGCTACAACGA ACGGGTCA	9954
1926	UAUAAAGA A UUUGGAGC	1923	GCTCCAAA GGCTAGCTACAACGA TCTTTATA	9955
1947	GUGGAGUU A CUCUCUUU	429	AAAGAGAG GGCTAGCTACAACGA AACTCCAC	9956
1967	GCCUUCUG A CUUCUUUC	1924	GAAAGAAG GGCTAGCTACAACGA CAGAAGGC	9957
1981	UUCCUUCU A UUCGAGAU	446	ATCTCGAA GGCTAGCTACAACGA AGAAGGAA	9958
1988	UAUUCGAG A UCUCCUCG	1925	CGAGGAGA GGCTAGCTACAACGA CTCGAATA	9959
1997	UCUCCUCG A CACCGCCU	1926	AGGCGGTG GGCTAGCTACAACGA CGAGGAGA	9960
1999	UCCUCGAC A CCGCCUCU	1172	AGAGGCGG GGCTAGCTACAACGA GTCGAGGA	9961
2015	UGCUCUGU A UCGGGGGG	454	CCCCCCGA GGCTAGCTACAACGA ACAGAGCA	9962
2040	UCUCCGGA A CAUUGUUC	1927	GAACAATG GGCTAGCTACAACGA TCCGGAGA	9963
2042	UCCGGAAC A UUGUUCAC	1183	GTGAACAA GGCTAGCTACAACGA GTTCCGGA	9964
2049	CAUUGUUC A CCUCACCA	1184	TGGTGAGG GGCTAGCTACAACGA GAACAATG	9965
2054	UUCACCUC A CCAUACGG	1187	CCGTATGG GGCTAGCTACAACGA GAGGTGAA	9966
2057	ACCUCACC A UACGGCAC	1189	GTGCCGTA GGCTAGCTACAACGA GGTGAGGT	9967
2059	CUCACCAU A CGGCACUC	464	GAGTGCCG GGCTAGCTACAACGA ATGGTGAG	9968
2064	CAUACGGC A CUCAGGCA	1190	TGCCTGAG GGCTAGCTACAACGA GCCGTATG	9969
2077	GGCAAGCU A UUCUGUGU	466	ACACAGAA GGCTAGCTACAACGA AGCTTGCC	9970
2098	GUGAGUUG A UGAAUCUA	1928	TAGATTCA GGCTAGCTACAACGA CAACTCAC	9971
2102	GUUGAUGA A UCUAGCCA	1929	TGGCTAGA GGCTAGCTACAACGA TCATCAAC	9972
2110	AUCUAGCC A CCUGGGUG	1198	CACCCAGG GGCTAGCTACAACGA GGCTAGAT	9973
2126	GGGAAGUA A UUUGGAAG	1930	CTTCCAAA GGCTAGCTACAACGA TACTTCCC	9974
2135	UUUGGAAG A UCCAGCAU	1931	ATGCTGGA GGCTAGCTACAACGA CTTCCAAA	9975
2142	GAUCCAGC A UCCAGGGA	1203	TCCCTGGA GGCTAGCTACAACGA GCTGGATC	9976
2151	UCCAGGGA A UUAGUAGU	1932	ACTACTAA GGCTAGCTACAACGA TCCCTGGA	9977
2165	AGUCAGCU A UGUCAACG	482	CGTTGACA GGCTAGCTACAACGA AGCTGACT	9978
2171	CUAUGUCA A CGUUAAUA	1933	TATTAACG GGCTAGCTACAACGA TGACATAG	9979
2177	CAACGUUA A UAUGGGCC	1934	GGCCCATA GGCTAGCTACAACGA TAACGTTG	9980
2179	ACGUUAAU A UGGGCCUA	486	TAGGCCCA GGCTAGCTACAACGA ATTAACGT	9981
2191	GCCUAAAA A UCAGACAA	1935	TTGTCTGA GGCTAGCTACAACGA TTTTAGGC	9982
2196	AAAAUCAG A CAACUAUU	1936	AATAGTTG GGCTAGCTACAACGA CTGATTTT	9983
2199	AUCAGACA A CUAUUGUG	1937	CACAATAG GGCTAGCTACAACGA TGTCTGAT	9984
2202	AGACAACU A UUGUGGUU	489	AACCACAA GGCTAGCTACAACGA AGTTGTCT	9985
2213	GUGGUUUC A CAUUUCCU	1214	AGGAAATG GGCTAGCTACAACGA GAAACCAC	9986
2215	GGUUUCAC A UUUCCUGU	1215	ACAGGAAA GGCTAGCTACAACGA GTGAAACC	9987
2227	CCUGUCUU A CUUUUGGG	499	CCCAAAAG GGCTAGCTACAACGA AAGACAGG	9988
2242	GGCGAGAA A CUGUUCUU	1938	AAGAACAG GGCTAGCTACAACGA TTCTCGCC	9989
2253	GUUCUUGA A UAUUUGGU	1939	ACCAAATA GGCTAGCTACAACGA TCAAGAAC	9990
2255	UCUUGAAU A UUUGGUGU	506	ACACCAAA GGCTAGCTACAACGA ATTCAAGA	9991
2278	GAGUGUGG A UUCGCACU	1940	AGTGCGAA GGCTAGCTACAACGA CCACACTC	9992
2284	GGAUUCGC A CUCCUCCU	1223	AGGAGGAG GGCTAGCTACAACGA GCGAATCC	9993
2295	CCUCCUGC A UAUAGACC	1229	GGTCTATA GGCTAGCTACAACGA GCAGGAGG	9994
2297	UCCUGCAU A UAGACCAC	517	GTGGTCTA GGCTAGCTACAACGA ATGCAGGA	9995
2301	GCAUAUAG A CCACCAAA	1941	TTTGGTGG GGCTAGCTACAACGA CTATATGC	9996
2304	UAUAGACC A CCAAAUGC	1231	GCATTTGG GGCTAGCTACAACGA GGTCTATA	9997
2309	ACCACCAA A UGCCCCUA	1942	TAGGGGCA GGCTAGCTACAACGA TTGGTGGT	9998
2317	AUGCCCCU A UCUUAUCA	519	TGATAAGA GGCTAGCTACAACGA AGGGGCAT	9999
2322	CCUAUCUU A UCAACACU	522	AGTGTTGA GGCTAGCTACAACGA AAGATAGG	10000
2326	UCUUAUCA A CACUUCCG	1943	CGGAAGTG GGCTAGCTACAACGA TGATAAGA	10001
2328	UUAUCAAC A CUUCCGGA	1240	TCCGGAAG GGCTAGCTACAACGA GTTGATAA	10002
2338	UUCCGGAA A CUACUGUU	1944	AACAGTAG GGCTAGCTACAACGA TTCCGGAA	10003

2341	CGGAAACU A CUGUUGUU	526	AACAACAG GGCTAGCTACAACGA AGTTTCCG	10004
2352	GUUGUUAG A CGAAGAGG	1945	CCTCTTCG GGCTAGCTACAACGA CTAACAAC	10005
2380	GAAGAAGA A CUCCCUCG	1946	CGAGGGAG GGCTAGCTACAACGA TCTTCTTC	10006
2397	CCUCGCAG A CGAAGGUC	1947	GACCTTCG GGCTAGCTACAACGA CTGCGAGG	10007
2409	AGGUCUCA A UCGCCGCG	1948	CGCGGCGA GGCTAGCTACAACGA TGAGACCT	10008
2427	CGCAGAAG A UCUCAAUC	1949	GATTGAGA GGCTAGCTACAACGA CTTCTGCG	10009
2433	AGAUCUCA A UCUCGGGA	1950	TCCCGAGA GGCTAGCTACAACGA TGAGATCT	10010
2442	UCUCGGGA A UCUCAAUG	1951	CATTGAGA GGCTAGCTACAACGA TCCCGAGA	10011
2448	GAAUCUCA A UGUUAGUA	1952	TACTAACA GGCTAGCTACAACGA TGAGATTC	10012
2456	AUGUUAGU A UUCCUUGG	547	CCAAGGAA GGCTAGCTACAACGA ACTAACAT	10013
2465	UUCCUUGG A CACAUAAG	1953	CTTATGTG GGCTAGCTACAACGA CCAAGGAA	10014
2467	CCUUGGAC A CAUAAGGU	1268	ACCTTATG GGCTAGCTACAACGA GTCCAAGG	10015
2469	UUGGACAC A UAAGGUGG	1269	CCACCTTA GGCTAGCTACAACGA GTGTCCAA	10016
2481	GGUGGGAA A CUUUACGG	1954	CCGTAAAG GGCTAGCTACAACGA TTCCCACC	10017
2486	GAAACUUU A CGGGGCUU	554	AAGCCCCG GGCTAGCTACAACGA AAAGTTTC	10018
2496	GGGGCUUU A UUCUUCUA	557	TAGAAGAA GGCTAGCTACAACGA AAAGCCCC	10019
2504	AUUCUUCU A CGGUACCU	562	AGGTACCG GGCTAGCTACAACGA AGAAGAAT	10020
2509	UCUACGGU A CCUUGCUU	563	AAGCAAGG GGCTAGCTACAACGA ACCGTAGA	10021
2520	UUGCUUUA A UCCUAAAU	1955	ATTTAGGA GGCTAGCTACAACGA TAAAGCAA	10022
2527	AAUCCUAA A UGGCAAAC	1956	GTTTGCCA GGCTAGCTACAACGA TTAGGATT	10023
2534	AAUGGCAA A CUCCUUCU	1957	AGAAGGAG GGCTAGCTACAACGA TTGCCATT	10024
2550	UUUUCCUG A CAUUCAUU	1958	AATGAATG GGCTAGCTACAACGA CAGGAAAA	10025
2552	UUCCUGAC A UUCAUUUG	1286	CAAATGAA GGCTAGCTACAACGA GTCAGGAA	10026
2556	UGACAUUC A UUUGCAGG	1287	CCTGCAAA GGCTAGCTACAACGA GAATGTCA	10027
2568	GCAGGAGG A CAUUGUUG	1959	CAACAATG GGCTAGCTACAACGA CCTCCTGC	10028
2570	AGGAGGAC A UUGUUGAU	1289	ATCAACAA GGCTAGCTACAACGA GTCCTCCT	10029
2577	CAUUGUUG A UAGAUGUA	1960	TACATCTA GGCTAGCTACAACGA CAACAATG	10030
2581	GUUGAUAG A UGUAAGCA	1961	TGCTTACA GGCTAGCTACAACGA CTATCAAC	10031
2590	UGUAAGCA A UUUGUGGG	1962	CCCACAAA GGCTAGCTACAACGA TGCTTACA	10032
2606	GGCCCCUU A CAGUAAAU	588	ATTTACTG GGCTAGCTACAACGA AAGGGGCC	10033
2613	UACAGUAA A UGAAAACA	1963	TGTTTTCA GGCTAGCTACAACGA TTACTGTA	10034
2619	AAAUGAAA A CAGGAGAC	1964	GTCTCCTG GGCTAGCTACAACGA TTTCATTT	10035
2626	AACAGGAG A CUUAAAUU	1965	AATTTAAG GGCTAGCTACAACGA CTCCTGTT	10036
2632	AGACUUAA A UUAACUAU	1966	ATAGTTAA GGCTAGCTACAACGA TTAAGTCT	10037
2636	UUAAAUUA A CUAUGCCU	1967	AGGCATAG GGCTAGCTACAACGA TAATTTAA	10038
2639	AAUUAACU A UGCCUGCU	594	AGCAGGCA GGCTAGCTACAACGA AGTTAATT	10039
2655	UAGGUUUU A UCCCAAUG	599	CATTGGGA GGCTAGCTACAACGA AAAACCTA	10040
2661	UUAUCCCA A UGUUACUA	1968	TAGTAACA GGCTAGCTACAACGA TGGGATAA	10041
2666	CCAAUGUU A CUAAAUAU	602	ATATTTAG GGCTAGCTACAACGA AACATTGG	10042
2671	GUUACUAA A UAUUUGCC	1969	GGCAAATA GGCTAGCTACAACGA TTAGTAAC	10043
2673	UACUAAAU A UUUGCCCU	604	AGGGCAAA GGCTAGCTACAACGA ATTTAGTA	10044
2685	GCCCUUAG A UAAAGGGA	1970	TCCCTTTA GGCTAGCTACAACGA CTAAGGGC	10045
2693	AUAAAGGG A UCAAACCG	1971	CGGTTTGA GGCTAGCTACAACGA CCCTTTAT	10046
2698	GGGAUCAA A CCGUAUUA	1972	TAATACGG GGCTAGCTACAACGA TTGATCCC	10047
2703	CAAACCGU A UUAUCCAG	611	CTGGATAA GGCTAGCTACAACGA ACGGTTTG	10048
2706	ACCGUAUU A UCCAGAGU	613	ACTCTGGA GGCTAGCTACAACGA AATACGGT	10049
2715	UCCAGAGU A UGUAGUUA	615	TAACTACA GGCTAGCTACAACGA ACTCTGGA	10050
2724	UGUAGUUA A UCAUUACU	1973	AGTAATGA GGCTAGCTACAACGA TAACTACA	10051
2727	AGUUAAUC A UUACUUCC	1313	GGAAGTAA GGCTAGCTACAACGA GATTAACT	10052
2730	UAAUCAUU A CUUCCAGA	621	TCTGGAAG GGCTAGCTACAACGA AATGATTA	10053
2738	ACUUCCAG A CGCGACAU	1974	ATGTCGCG GGCTAGCTACAACGA CTGGAAGT	10054

2743	CAGACGCG A CAUUAUUU	1975	AAATAATG GGCTAGCTACAACGA CGCGTCTG	10055
2745	GACGCGAC A UUAUUUAC	1317	GTAAATAA GGCTAGCTACAACGA GTCGCGTC	10056
2748	GCGACAUU A UUUACACA	625	TGTGTAAA GGCTAGCTACAACGA AATGTCGC	10057
2752	CAUUAUUU A CACACUCU	628	AGAGTGTG GGCTAGCTACAACGA AAATAATG	10058
2754	UUAUUUAC A CACUCUUU	1318	AAAGAGTG GGCTAGCTACAACGA GTAAATAA	10059
2756	AUUUACAC A CUCUUUGG	1319	CCAAAGAG GGCTAGCTACAACGA GTGTAAAT	10060
2774	AGGCGGGG A UCUUAUAU	1976	ATATAAGA GGCTAGCTACAACGA CCCCGCCT	10061
2779	GGGAUCUU A UAUAAAAG	634	CTTTTATA GGCTAGCTACAACGA AAGATCCC	10062
2781	GAUCUUAU A UAAAAGAG	635	CTCTTTTA GGCTAGCTACAACGA ATAAGATC	10063
2795	GAGAGUCC A CACGUAGC	1324	GCTACGTG GGCTAGCTACAACGA GGACTCTC	10064
2797	GAGUCCAC A CGUAGCGC	1325	GCGCTACG GGCTAGCTACAACGA GTGGACTC	10065
2809	AGCGCCUC A UUUUGCGG	1328	CCGCAAAA GGCTAGCTACAACGA GAGGCGCT	10066
2821	UGCGGGUC A CCAUAUUC	1329	GAATATGG GGCTAGCTACAACGA GACCCGCA	10067
2824	GGGUCACC A UAUUCUUG	1331	CAAGAATA GGCTAGCTACAACGA GGTGACCC	10068
2826	GUCACCAU A UUCUUGGG	644	CCCAAGAA GGCTAGCTACAACGA ATGGTGAC	10069
2836	UCUUGGGA A CAAGAUCU	1977	AGATCTTG GGCTAGCTACAACGA TCCCAAGA	10070
2841	GGAACAAG A UCUACAGC	1978	GCTGTAGA GGCTAGCTACAACGA CTTGTTCC	10071
2845	CAAGAUCU A CAGCAUGG	649	CCATGCTG GGCTAGCTACAACGA AGATCTTG	10072
2850	UCUACAGC A UGGGAGGU	1336	ACCTCCCA GGCTAGCTACAACGA GCTGTAGA	10073
2870	UCUUCCAA A CCUCGAAA	1979	TTTCGAGG GGCTAGCTACAACGA TTGGAAGA	10074
2883	GAAAAGGC A UGGGGACA	1342	TGTCCCCA GGCTAGCTACAACGA GCCTTTTC	10075
2889	GCAUGGGG A CAAAUCUU	1980	AAGATTTG GGCTAGCTACAACGA CCCCATGC	10076
2893	GGGGACAA A UCUUUCUG	1981	CAGAAAGA GGCTAGCTACAACGA TTGTCCCC	10077
2908	UGUCCCCA A UCCCCUGG	1982	CCAGGGGA GGCTAGCTACAACGA TGGGGACA	10078
2918	CCCCUGGG A UUCUUCCC	1983	GGGAAGAA GGCTAGCTACAACGA CCCAGGGG	10079
2929	CUUCCCCG A UCAUCAGU	1984	ACTGATGA GGCTAGCTACAACGA CGGGGAAG	10080
2932	CCCCGAUC A UCAGUUGG	1358	CCAACTGA GGCTAGCTACAACGA GATCGGGG	10081
2941	UCAGUUGG A CCCUGCAU	1985	ATGCAGGG GGCTAGCTACAACGA CCAACTGA	10082
2948	GACCCUGC A UUCAAAGC	1363	GCTTTGAA GGCTAGCTACAACGA GCAGGGTC	10083
2959	CAAAGCCA A CUCAGUAA	1986	TTACTGAG GGCTAGCTACAACGA TGGCTTTG	10084
2968	CUCAGUAA A UCCAGAUU	1987	AATCTGGA GGCTAGCTACAACGA TTACTGAG	10085
2974	AAAUCCAG A UUGGGACC	1988	GGTCCCAA GGCTAGCTACAACGA CTGGATTT	10086
2980	AGAUUGGG A CCUCAACC	1989	GGTTGAGG GGCTAGCTACAACGA CCCAATCT	10087
2986	GGACCUCA A CCCGCACA	1990	TGTGCGGG GGCTAGCTACAACGA TGAGGTCC	10088
2998	GCACAAGG A CAACUGGC	1991	GCCAGTTG GGCTAGCTACAACGA CCTTGTGC	10089
3001	CAAGGACA A CUGGCCGG	1992	CCGGCCAG GGCTAGCTACAACGA TGTCCTTG	10090
3010	CUGGCCGG A CGCCAACA	1993	TGTTGGCG GGCTAGCTACAACGA CCGGCCAG	10091
3016	GGACGCCA A CAAGGUGG	1994	CCACCTTG GGCTAGCTACAACGA TGGCGTCC	10092
3035	GUGGGAGC A UUCGGGCC	1384	GGCCCGAA GGCTAGCTACAACGA GCTCCCAC	10093
3051	CAGGGUUC A CCCCUCCC	1387	GGGAGGGG GGCTAGCTACAACGA GAACCCTG	10094
3061	CCCUCCCC A UGGGGGAC	1395	GTCCCCCA GGCTAGCTACAACGA GGGGAGGG	10095
3068	CAUGGGGG A CUGUUGGG	1995	CCCAACAG GGCTAGCTACAACGA CCCCCATG	10096
3088	GAGCCCUC A CGCUCAGG	1400	CCTGAGCG GGCTAGCTACAACGA GAGGGCTC	10097
3101	CAGGGCCU A CUCACAAC	683	GTTGTGAG GGCTAGCTACAACGA AGGCCCTG	10098
3105	GCCUACUC A CAACUGUG	1406	CACAGTTG GGCTAGCTACAACGA GAGTAGGC	10099
3108	UACUCACA A CUGUGCCA	1996	TGGCACAG GGCTAGCTACAACGA TGTGAGTA	10100
3138	CUGCCUCC A CCAAUCGG	1422	CCGATTGG GGCTAGCTACAACGA GGAGGCAG	10101
3142	CUCCACCA A UCGGCAGU	1997	ACTGCCGA GGCTAGCTACAACGA TGGTGGAG	10102
3165	GGCAGCCU A CUCCCUUA	691	TAAGGGAG GGCTAGCTACAACGA AGGCTGCC	10103
3173	ACUCCCUU A UCUCCACC	694	GGTGGAGA GGCTAGCTACAACGA AAGGGAGT	10104
3179	UUAUCUCC A CCUCUAAG	1436	CTTAGAGG GGCTAGCTACAACGA GGAGATAA	10105

3190	UCUAAGGG A CACUCAUC	1998	GATGAGTG GGCTAGCTACAACGA CCCTTAGA	10106
3192	UAAGGGAC A CUCAUCCU	1440	AGGATGAG GGCTAGCTACAACGA GTCCCTTA	10107
3196	GGACACUC A UCCUCAGG	1442	CCTGAGGA GGCTAGCTACAACGA GAGTGTCC	10108
3207	CUCAGGCC A UGCAGUGG	1447	CCACTGCA GGCTAGCTACAACGA GGCCTGAG	10109

Input Sequence = AF100308. Cut Site = YG/M or UG/U.
Stem Length = 8 . Core Sequence = GGCTAGCTACAACGA
AF100308 (Hepatitis B virus strain 2-18, 3215 bp)

TABLE X: HUMAN HBV AMBERZYME AND SUBSTRATE SEQUENCE

Pos	Substrate	Seq ID	Amberzyme	Seq ID
61	ACUUUCCU G CUGGUGGC	1448	GCCACCAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGGAAAGU	10110
87	GGAACAGU G AGCCCUGC	1449	GCAGGGCU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG ACUGUUCC	10111
94	UGAGCCCU G CUCAGAAU	1450	AUUCUGAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGGGCUCA	10112
112	CUGUCUCU G CCAUAUCG	1451	CGAUAUGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGAGACAG	10113
132	AUCUUAUC G AAGACUGG	1452	CCAGUCUU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GAUAAGAU	10114
153	CCUGUACC G AACAUGGA	1453	UCCAUGUU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGUACAGG	10115
169	AGAACAUC G CAUCAGGA	1454	UCCUGAUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GAUGUUCU	10116
192	GGACCCCU G CUCGUGUU	1455	AACACGAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGGGGUCC	10117
222	UUCUUGUU G ACAAAAAU	1456	AUUUUUGU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AACAAGAA	10118
315	CAAAAUUC G CAGUCCCA	1457	UGGGACUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GAAUUUUG	10119
374	UGGUUAUC G CUGGAUGU	1458	ACAUCCAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GAUAACCA	10120
387	AUGUGUCU G CGGCGUUU	1459	AAACGCCG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGACACAU	10121
410	CUUCCUCU G CAUCCUGC	1460	GCAGGAUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGAGGAAG	10122
417	UGCAUCCU G CUGCUAUG	1461	CAUAGCAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGGAUGCA	10123
420	AUCCUGCU G CUAUGCCU	1462	AGGCAUAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGCAGGAU	10124
425	GCUGCUAU G CCUCAUCU	1463	AGAUGAGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AUAGCAGC	10125
468	GGUAUGUU G CCCGUUUG	1464	CAAACGGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AACAUACC	10126
518	CGGACCAU G CAAAACCU	1465	AGGUUUUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AUGGUCCG	10127
527	CAAAACCU G CACAACUC	1466	GAGUUGUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGGUUUUG	10128
538	CAACUCCU G CUCAAGGA	1467	UCCUUGAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGGAGUUG	10129
569	CUCAUGUU G CUGUACAA	1468	UUGUACAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AACAUGAG	10130
596	CGGAAACU G CACCUGUA	1469	UACAGGUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGUUUCCG	10131
631	GGGCUUUC G CAAAAUAC	1470	GUAUUUUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GAAAGCCC	10132
687	UUACUAGU G CCAUUUGU	1471	ACAAAUGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG ACUAGUAA	10133
747	AUAUGGAU G AUGUGGUU	1472	AACCACAU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AUCCAUAU	10134
783	AACAUCUU G AGUCCCUU	1473	AAGGGACU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AAGAUGUU	10135
795	CCCUUUAU G CCGCUGUU	1474	AACAGCGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AUAAAGGG	10136
798	UUUAUGCC G CUGUUACC	1475	GGUAACAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGCAUAAA	10137
911	GGCACAUU G CCACAGGA	1476	UCCUGUGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AAUGUGCC	10138
978	GGCCUAUU G AUUGGAAA	1477	UUUCCAAU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AAUAGGCC	10139
997	AUGUCAAC G AAUUGUGG	1478	CCACAAUU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GUUGACAU	10140
1020	UGGGGUUU G CCGCCCCU	1479	AGGGGCGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AAACCCCA	10141
1023	GGUUUGCC G CCCCUUUC	1480	GAAAGGGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGCAAACC	10142

1034	CCUUUCAC G CAAUGUGG	1481	CCACAUUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GUGAAAGG	10143
1050	GAUAUUCU G CUUUAAUG	1482	CAUUAAAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGAAUAUC	10144
1058	GCUUUAAU G CCUUUAUA	1483	UAUAAAGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AUUAAAGC	10145
1068	CUUUAUAU G CAUGCAUA	1484	UAUGCAUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AUAUAAAG	10146
1072	AUAUGCAU G CAUACAAG	1485	CUUGUAUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AUGCAUAU	10147
1103	ACUUUCUC G CCAACUUA	1486	UAAGUUGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GAGAAAGU	10148
1139	CAGUAUGU G AACCUUUA	1487	UAAAGGUU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG ACAUACUG	10149
1155	ACCCCGUU G CUCGGCAA	1488	UUGCCGAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AACGGGGU	10150
1177	UGGUCUAU G CCAAGUGU	1489	ACACUUGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AUAGACCA	10151
1188	AAGUGUUU G CUGACGCA	1490	UGCGUCAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AAACACUU	10152
1191	UGUUUGCU G ACGCAACC	1491	GGUUGCGU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGCAAACA	10153
1194	UUGCUGAC G CAACCCCC	1492	GGGGGUUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GUCAGCAA	10154
1234	CCAUCAGC G CAUGCGUG	1493	CACGCAUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GCUGAUGG	10155
1238	CAGCGCAU G CGUGGAAC	1494	GUUCCACG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AUGCGCUG	10156
1262	UCUCCUCU G CCGAUCCA	1495	UGGAUCGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGAGGAGA	10157
1265	CCUCUGCC G AUCCAUAC	1496	GUAUGGAU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGCAGAGG	10158
1275	UCCAUACC G CGGAACUC	1497	GAGUUCCG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGUAUGGA	10159
1290	UCCUAGCC G CUUGUUUU	1498	AAAACAAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGCUAGGA	10160
1299	CUUGUUUU G CUCGCAGC	1499	GCUGCGAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AAAACAAG	10161
1303	UUUUGCUC G CAGCAGGU	1500	ACCUGCUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GAGCAAAA	10162
1335	UCGGGACU G ACAAUUCU	1501	AGAAUUGU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGUCCCGA	10163
1349	UCUGUCGU G CUCUCCCG	1502	CGGGAGAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG ACGACAGA	10164
1357	GCUCUCCC G CAAAUAUA	1503	UAUAUUUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGGAGAGC	10165
1382	CCAUGGCU G CUAGGCUG	1504	CAGCCUAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGCCAUGG	10166
1392	UAGGCUGU G CUGCCAAC	1505	GUUGGCAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG ACAGCCUA	10167
1395	GCUGUGCU G CCAACUGG	1506	CCAGUUGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGCACAGC	10168
1411	GAUCCUAC G CGGGACGU	1507	ACGUCCCG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GUAGGAUC	10169
1442	CCGUCGGC G CUGAAUCC	1508	GGAUUCAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GCCGACGG	10170
1445	UCGGCGCU G AAUCCCGC	1509	GCGGGAUU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG AGCGCCGA	10171
1452	UGAAUCCC G CGGACGAC	1510	GUCGUCCG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGGAUUCA	10172
1458	CCGCGGAC G ACCCCUCC	1511	GGAGGGGU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GUCCGCGG	10173
1474	CCGGGGCC G CUUGGGGC	1512	GCCCCAAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGCCCCGG	10174
1489	GCUCUACC G CCCGCUUC	1513	GAAGCGGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGUAGAGC	10175
1493	UACCGCCC G CUUCUCCG	1514	CGGAGAAG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGGCGGUA	10176
1501	GCUUCUCC G CCUAUUGU	1515	ACAAUAGG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGAGAAGC	10177
1513	AUUGUACC G ACCGUCCA	1516	UGGACGGU GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GGUACAAU	10178
1528	CACGGGGC G CACCUCUC	1517	GAGAGGUG GGAGGAAACUCC CU UCAAGGACAUCGUCCGGG GCCCCGUG	10179

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 193

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 193

NOTE: For additional volumes, please contact the Canadian Patent Office

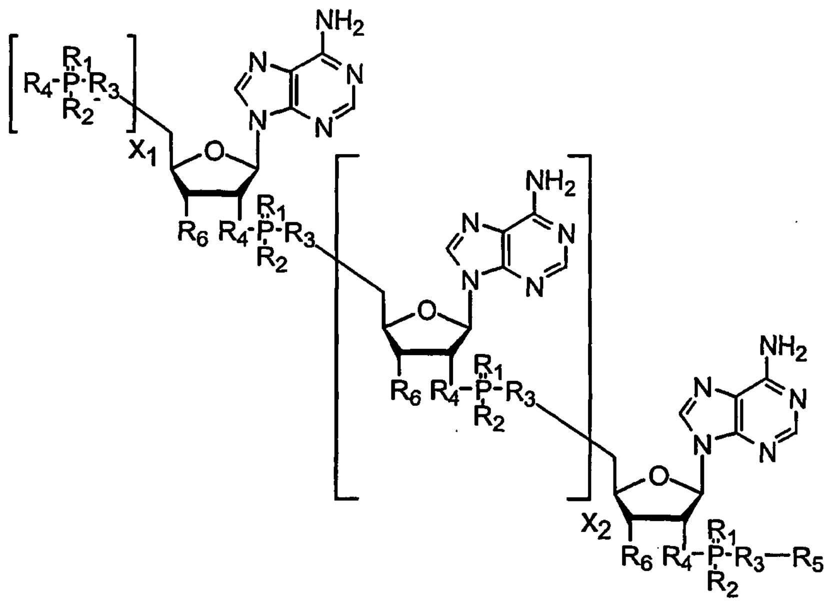
NOM DU FICHIER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

CLAIMS

What we claim is:

1. A compound having Formula I:



wherein X_1 is an integer selected from the group consisting of 1, 2, and 3; X_2 is an integer greater than or equal to 1; R_6 is independantly selected from the group consisting of H, OH, NH_2, O NH_2, alkyl, S-alkyl, O-alkyl, O-alkyl-S-alkyl, O-alkoxyalkyl, allyl, O-allyl, and fluoro; each R_1 and R_2 are independantly selected from the group consisting of O and S; each R_3 and R_4 are independantly selected from the group consisting of O, N, and S; and R_5 is selected from the group consisting of alkyl, alkylamine, oligonucleotide having any of SEQ ID NOS. 11343-16182, oligonucleotide having a sequence complementary to any of SEQ ID NOS. 2594-7433, and abasic moiety.

2. The compound of claim 1, wherein said oligonucleotide having a sequence complementary to any of SEQ ID NOS. 2594-7433 is an enzymatic nucleic acid molecule.

3. The compound of claim 1, wherein said oligonucleotide having a sequence complementary to any of SEQ ID NOS. 2594-7433 is an antisense nucleic acid molecule.

4. The compound of claim 2, wherein said enzymatic nucleic acid molecule is selected from the group consisting of Hammerhead, Inozyme, G-cleaver, DNAzyme, Amberzyme, and Zinzyme motifs.

5. The compound of claim 2, wherein said Inozyme enzymatic nucleic acid molecule comprises a stem II region of length greater than or equal to 2 base pairs.

6. The compound of claim 2, wherein said enzymatic nucleic acid comprises between 12 and 100 bases complementary to an RNA derived from HCV.

7. The compound of claim 2, wherein said enzymatic nucleic acid comprises between 14 and 24 bases complementary to an RNA derived from HCV.

8. The compound of claim 3, wherein said antisense nucleic acid comprises between 12 and 100 bases complementary to an RNA derived from HCV.

9. The compound of claim 3, wherein said antisense nucleic acid comprises between 14 and 24 bases complementary to an RNA derived from HCV.

10. A composition comprising the compound of claim 1and a pharmaceutically acceptable carrier.

11. A mammalian cell comprising a compound of claim 1.

12. The mammalian cell of claim 11, wherein said mammalian cell is a human cell.

13. A method for treatment of cirrhosis, liver failure, hepatocellular carcinoma, or a condition associated with HCV infection comprising the step of administering to a patient a compound of claim 1 under conditions suitable for said treatment.

14. The method of claim 13 further comprising the use of one or more drug therapies under conditions suitable for said treatment.

15. A method for inhibiting HCV replication in a mammalian cell comprising the step of administering to said cell the compound of claim 1 under conditions suitable for said inhibition.

16. A method of cleaving a separate RNA molecule comprising contacting the compound of claim 1 with said separate RNA molecule under conditions suitable for the cleavage of said separate RNA molecule.

17. The method of claim 16, wherein said cleavage is carried out in the presence of a divalent cation.

18. The method of claim 17, wherein said divalent cation is Mg^{2+}.

19. The method of claim 16, wherein said cleavage is carried out in the presence of a protein nuclease.

20. The method of claim 19, wherein said protein nuclease is an RNAse L.

21. The compound of claim 1, wherein said compound is chemically synthesized.

22. The compound of claim 1, wherein said oligonucleotide comprises at least one 2'-sugar modification.

23. The compound of claim 1, wherein said oligonucleotide comprises at least one nucleic acid base modification.

24. The compound of claim 1, wherein said oligonucleotide comprises at least one phosphate modification.

25. The method of claim 14, wherein said drug therapy is the administration of type I interferon.

26. The method of claim 25, wherein said type I interferon and the compound of claim 1 are administered simultaneously.

27. The method of claim 25, wherein said type I interferon and the compound of claim 1 are administered separately.

28. The method of claim 25, wherein said type I interferon is selected from the group consisting of interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon,

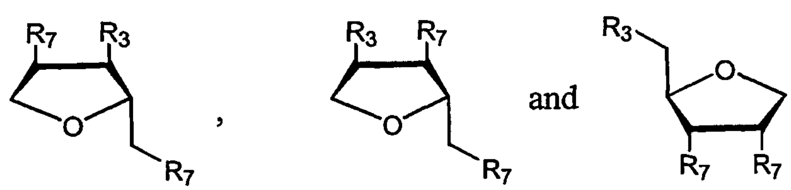
polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, and polyethylene glycol consensus interferon.

29. The method of claim 14, wherein R_5 in said compound is selected from the group consisting of alkyl, alkylamine and abasic moiety and said drug therapy comprises treatment with an enzymatic nucleic acid molecule which is targeted against HCV replication.

30. The method of claim 14, wherein R_5 in said compound is selected from the group consisting of alkyl, alkylamine and abasic moiety and said drug therapy comprises treatment with an antisense nucleic acid molecule which is targeted against HCV replication.

31. A composition comprising type I interferon and the compound of claim 1 and a pharmaceutically acceptable carrier.

32. The compound of claim 1, wherein said abasic moiety is selected from the group consisting of:



wherein R_3 is selected from the group consisting of S, N, or O and R_7 is independently selected from the group consisting of H, OH, NH2, O-NH2, alkyl, S-alkyl, O-alkyl, O-alkyl-S-alkyl, O-alkoxyalkyl, allyl, O-allyl, fluoro, oligonucleotide, alkyl, alkylamine and abasic moiety.

33. An enzymatic nucleic acid molecule that specifically cleaves RNA derived from hepatitis B virus (HBV), wherein said enzymatic nucleic acid molecule comprises sequence defined as Seq. ID No. 6346.

34. A method of administering to a cell an enzymatic nucleic acid molecule of claim 33 comprising contacting said cell with the enzymatic nucleic acid molecule under conditions suitable for said administration.

35. The method of claim 34, further comprising the administration of one or more other therapeutic compounds.

36. The method of claim 35, wherein said other therapeutic compound is type I interferon.

37. The method of claim 35, wherein said other therapeutic compound is 3TC® (Lamivudine).

38. The method of claim 35, wherein said other therapeutic compound and the enzymatic nucleic acid molecule are administered simultaneously.

39. The method of claim 35, wherein said other therapeutic compound and enzymatic nucleic acid molecule are administered separately.

40. The method of claim 36, wherein said type I interferon is selected from the group consisting of interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, and polyethylene glycol consensus interferon.

41. The method of claim 34 or claim 35, wherein said cell is a mammalian cell.

42. The method of claim 41, wherein said cell is a human cell.

43. The method of claim 41, wherein said administration is in the presence of a delivery reagent.

44. The method of claim 43, wherein said delivery reagent is a lipid.

45. The method of claim 44, wherein said lipid is a cationic lipid or a phospholipid.

46. The method of claim 43, wherein said delivery reagent is a liposome.

47. A nucleic acid molecule that specifically binds the hepatitis B virus (HBV) reverse transcriptase primer, wherein said nucleic acid molecule comprises the sequence $(UUCA)_n$, wherein n is an integer from 1 to 10.

48. A nucleic acid molecule that specifically binds the hepatitis B virus (HBV) reverse transcriptase primer, wherein said nucleic acid molecule is a sequence comprising any of Seq. ID Nos: 11216-11262, 11264, 11266, 11268, 11270, 11272, 11274, 11276, 11278, 11280, 11282, 11284, 11286, 11288, 11290 and 11292.

49. A nucleic acid molecule that specifically binds to the Enhancer I sequence of HBV DNA.

50. A nucleic acid molecule of claim 49 wherein said nucleic acid molecule comprises any of SEQ ID Nos: 11327, 11330, 11332, 11334, 11335, 11338, 11340 and 11342.

51. A method of administering to a cell a nucleic acid molecule of any of claims 47-50 comprising contacting said cell with the nucleic acid decoy molecule under conditions suitable for said administration.

52. The method of claim 51, further comprising administering one or more other therapeutic compounds.

53. The method of claim 52, wherein said other therapeutic compound is type I interferon.

54. The method of claim 52, wherein said other therapeutic compound is 3TC® (Lamivudine).

55. The method of claim 52, wherein said other therapeutic compound and the nucleic acid molecule are administered simultaneously.

56. The method of claim 52, wherein said other therapeutic compound and the nucleic acid molecule are administered separately.

57. The method of claim 53, wherein said type I interferon is selected from the group consisting of interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, and polyethylene glycol consensus interferon.

58. The nucleic acid molecule of any of claims 47-50, wherein said nucleic acid molecule comprises a nucleic acid backbone modification.

59. The nucleic acid molecule of any of claims 47-50, wherein said nucleic acid molecule comprises a nucleic acid sugar modification.

60. The nucleic acid molecule of any of claims 47-50, wherein said nucleic acid decoy molecule comprises a nucleic acid base modification.

61. The method of claim 51 or claim 52, wherein said cell is a mammalian cell.

62. The method of claim 61, wherein said cell is a human cell.

63. The method of claim 61, wherein said administration is in the presence of a delivery reagent.

64. The method of claim 63, wherein said delivery reagent is a lipid.

65. The method of claim 64, wherein said lipid is a cationic lipid or a phospholipid.

66. The method of claim 63 wherein said delivery reagent is a liposome.

67. The nucleic acid molecule of claim 47, wherein said nucleic acid molecule is a decoy nucleic acid molecule.

68. The nucleic acid molecule of claim 47, wherein said nucleic acid molecule is an aptamer nucleic acid molecule.

69. The nucleic acid molecule of claim 49, wherein said Enhancer I sequence comprises a Hepatocyte Nuclear Factor 3 and/or Hepatocyte Nuclear Factor 4 binding sequence.

70. A mouse implanted with HepG2.2.15 cells, wherein said mouse sustains the propagation of HEPG2.2.15 cells and HBV production.

71. The mouse of claim 70, wherein said mouse has been infected with HBV for at least one week.

72. The mouse of claim 70, wherein said mouse has been infected with HCV for at least four weeks.

73. The mouse of claim 70, wherein said mouse has been infected with HBV for at least eight weeks.

74. The mouse of claim 70, wherein said mouse is an immuno compromised mouse.

75. The mouse of claim 74, wherein said mouse is a nu/nu mouse.

76. The mouse of claim 74, wherein said mouse is a scid/scid mouse.

77. A method of producing a mouse according to claim 70, comprising injecting HepG2.2.15 cells into said mouse under conditions suitable for the propagation of the HepG2.2.15 cells in said mouse.

78. The method of claim 77, wherein said mouse is a nu/nu mouse.

79. The method of claim 77, wherein said mouse is a scid/scid mouse.

80. The method of claim 77, wherein said injection is subcutaneous injection.

81. The method of claim 77, wherein said HepG2.2.15 cells are suspended in Dulbecco's PBS solution including calcium and magnesium.

82. A method of screening a therapeutic compound for activity against HBV comprising administering said therapeutic compound to a mouse of claim 70 and monitoring said mouse for the effects of said therapeutic compound on levels of HBV DNA.

83. The method of claim 70, wherein said therapeutic compound is a nucleic acid molecule, administered alone or in combination with another therapeutic compound or treatment.

84. The method of claim 83, wherein said nucleic acid molecule is an enzymatic nucleic acid molecule.

85. The method of claim 83, wherein said nucleic acid molecule is an antisense nucleic acid molecule.

86. The method of claim 83, wherein said other treatment is antiviral therapy.

87. The method of claim 86, wherein said antiviral therapy is treatment with 3TC® (Lamivudine).

88. The method of claim 86, wherein said antiviral therapy is treatment with interferon.

89. The method of claim 88, wherein said interferon is selected from the group consisting of consensus interferon, type I interferon, interferon alpha, interferon beta, consensus

interferon, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b and polyethylene glycol consensus interferon.

90. An immunocompromised non-human mammal implanted with HepG2.2.15 cells, wherein said non-human mammal is susceptible to HBV infection and capable of sustaining HBV DNA expression.

91. The mammal of claim 90, wherein said non-human mammal has been infected with HBV for at least one week.

92. The mammal of claim 90, wherein said non-human mammal has been infected with HCV for at least four weeks.

93. The mammal of claim 90, wherein said non-human mammal has been infected with HBV for at least eight weeks.

94. The mammal of claim 90, wherein said non-human mammal is a nu/nu mammal.

95. The mammal of claim 90, wherein said non-human mammal is a scid/scid mammal.

96. A method of producing a non-human mammal according to claim 90, comprising injecting HepG2.2.15 cells into said non-human mammal under conditions suitable for the propagation of the HepG2.2.15 cells in said non-human.

97. The method of claim 96, wherein said non-human mammal is a nu/nu mammal.

98. The method of claim 96, wherein said non-human mammal is a scid mammal.

99. The method of claim 96, wherein said injection is subcutaneous injection.

100. The method of claim 96, wherein said HepG2.2.15 cells are suspended in Delbecco's PBS solution including calcium and magnesium.

101. A method of screening a therapeutic compound for activity against HBV, comprising administering said therapeutic compound to a non-human mammal of claim 90 and monitoring said mammal for the effects of said therapeutic compound on levels of HBV DNA.

102. The method of claim 101, wherein said therapeutic compound is a nucleic acid molecule administered alone or in combination with another therapeutic compound or treatment.

103.The method of claim 102, wherein said nucleic acid molecule is an enzymatic nucleic acid molecule.

104.The method of claim 102, wherein said nucleic acid molecule is an antisense nucleic acid molecule.

105.The method of claim 102, wherein said other treatment is antiviral therapy.

106.The method of claim 105, wherein said antiviral therapy is treatment with 3TC® (Lamivudine).

107.The method of claim 105, wherein said antiviral therapy is treatment with interferon.

108.The method of claim 107, wherein said interferon is selected from the group consisting of consensus interferon, type I interferon, interferon alpha, interferon beta, consensus interferon, polyethylene glycol interferon, polyethylene glycol interferon alpha 2a, polyethylene glycol interferon alpha 2b, and polyethylene glycol consensus interferon.

Figure 1: Ribozyme Motifs

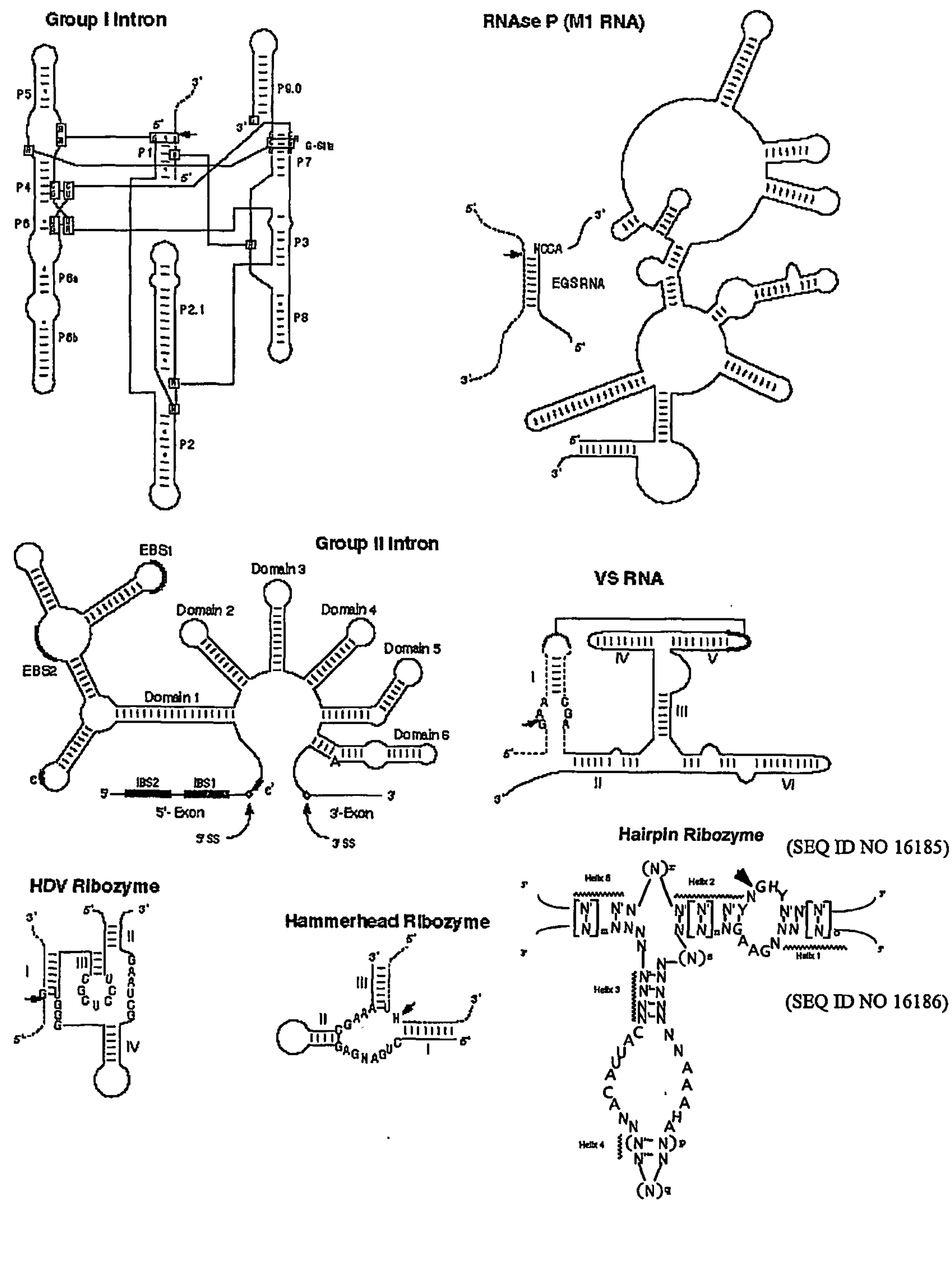


Figure 2: Examples of Nuclease Stable Ribozyme Motifs

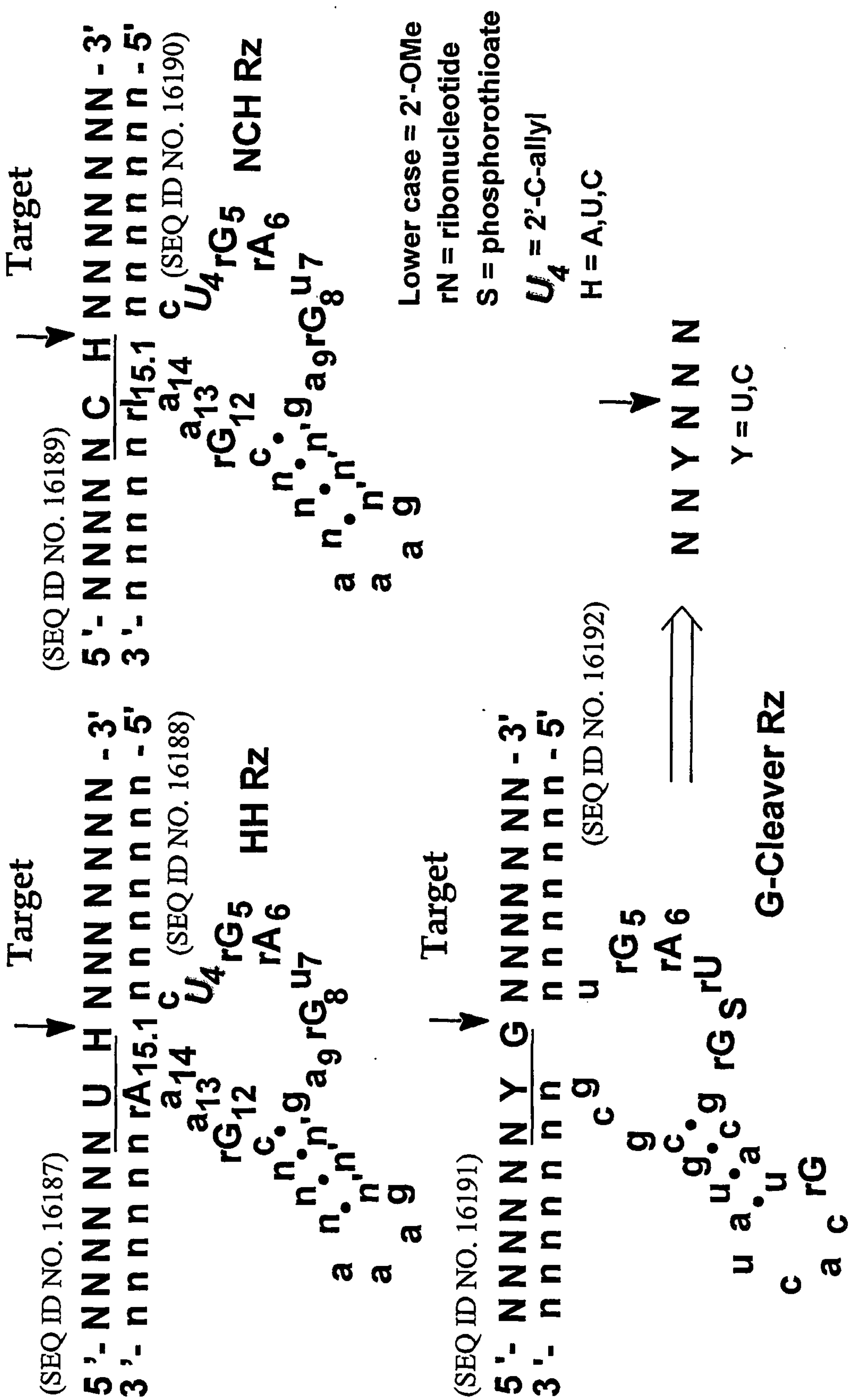


Figure 3: 2'-O-Me substituted Amberzyme Enzymatic Nucleic Acid Motif

U,C = 2'-NH$_2$-U,C
Lower case = 2'-O-Me
Uppercase = Ribo

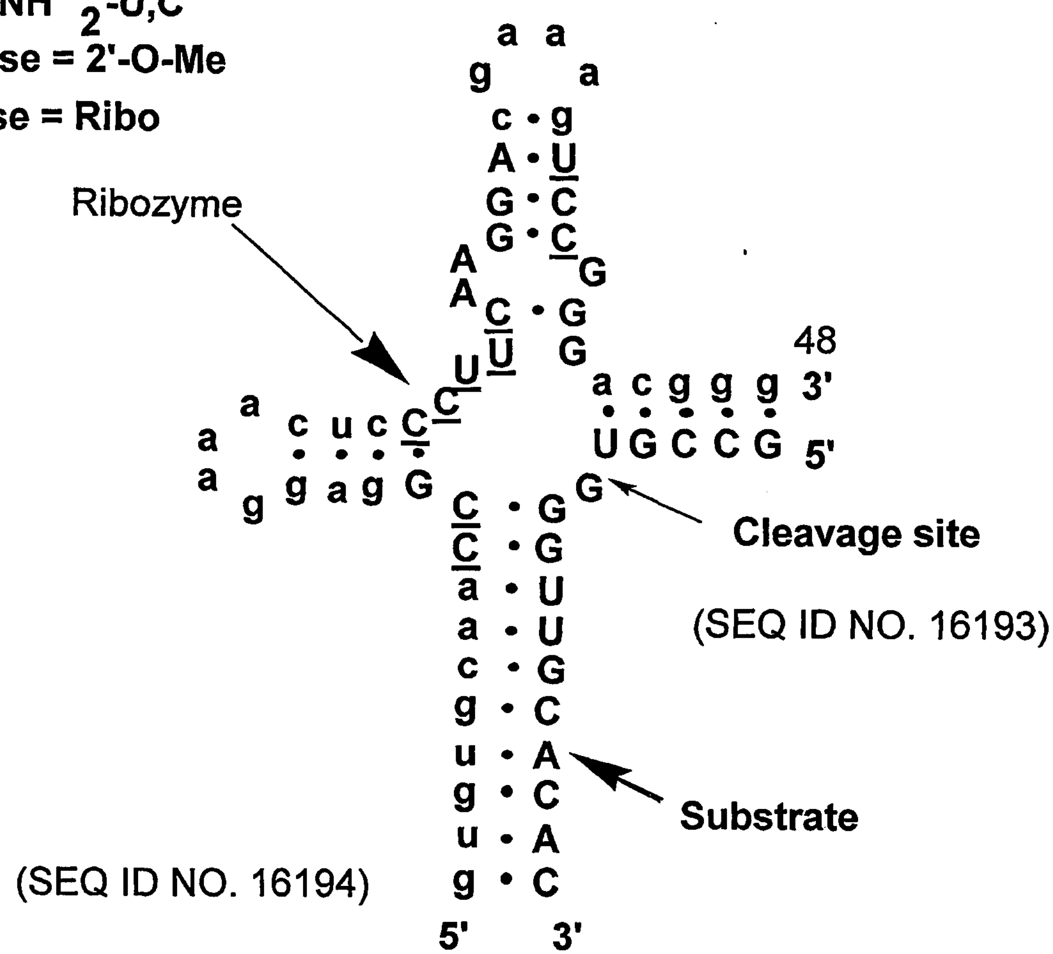
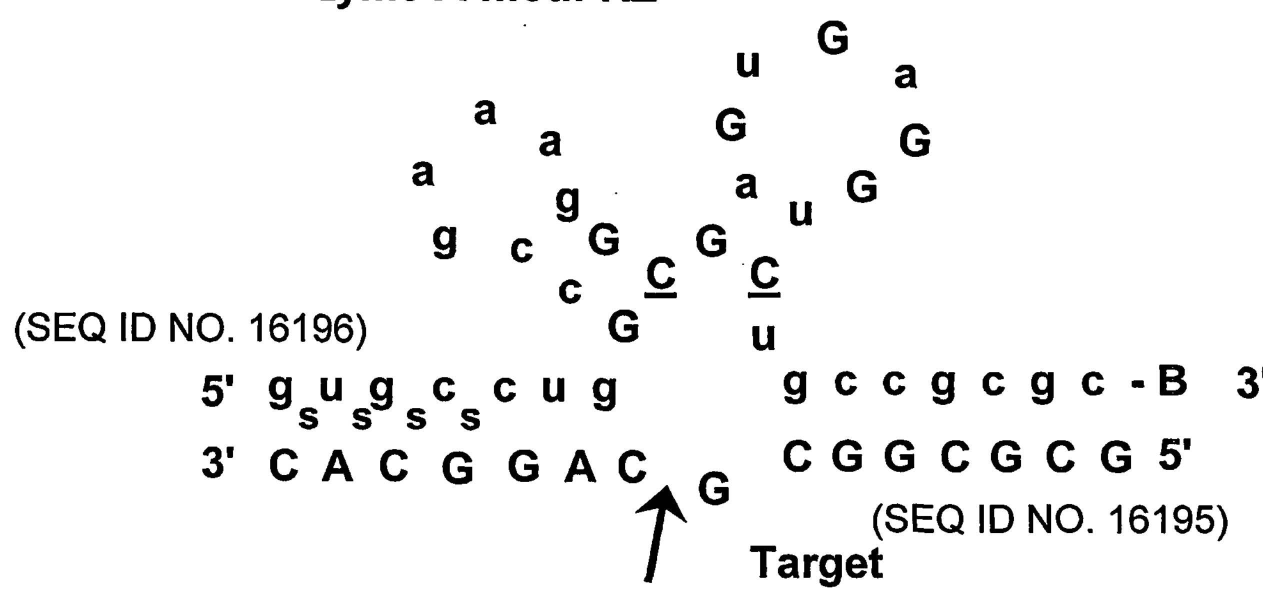


Figure 4: Stabilized Zinzyme Ribozyme Motif

Zinzyme A-motif RZ



Legend

Uppercase indicates natural ribo residues

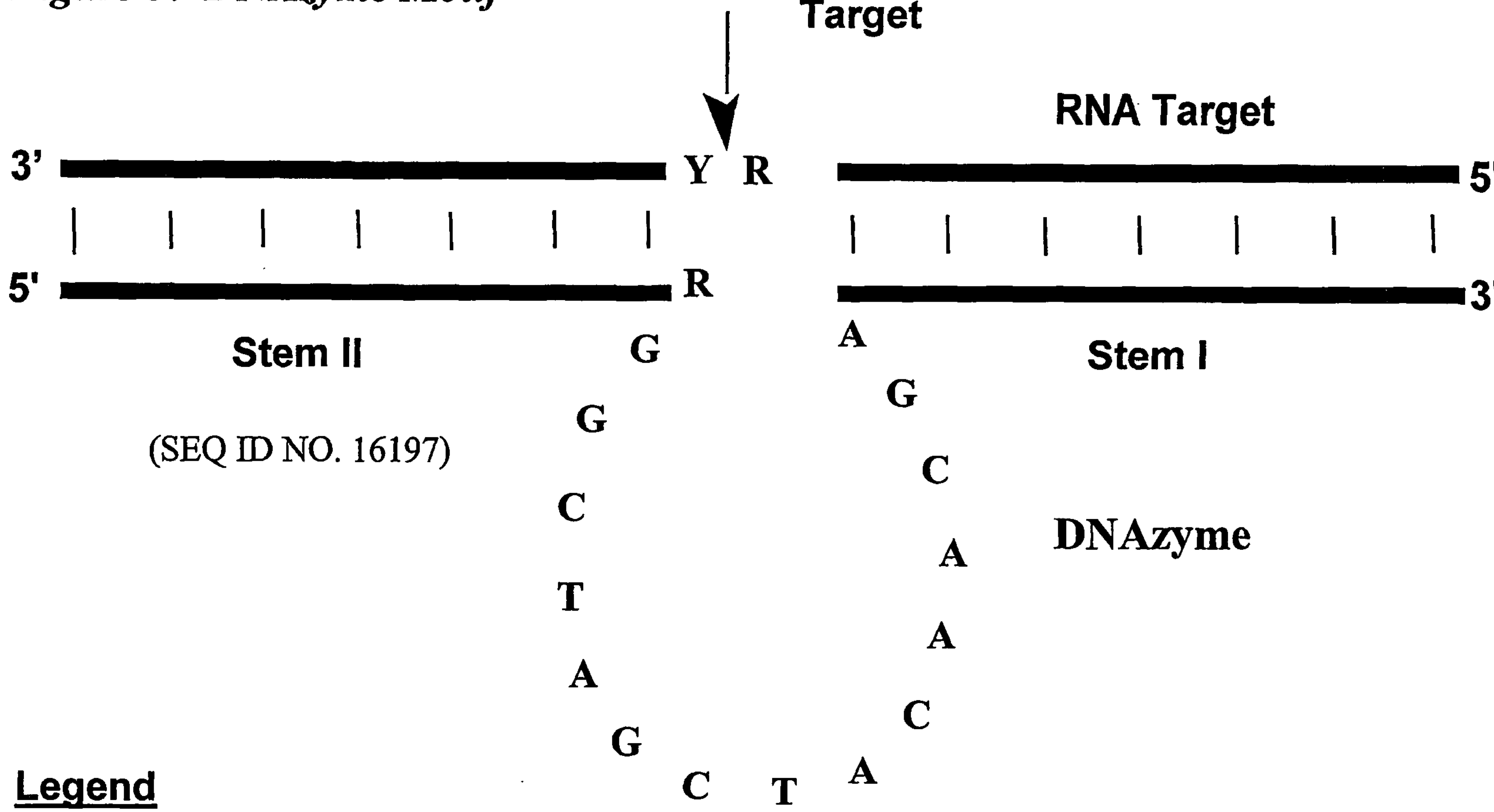
C indicates 2'- d-NH_2-C

Lowercase: 2'-O-Me

Subscript $_s$ indicates phosphothioate linkage

B: 3'- 3' abasic moiety

Figure 5: DNAzyme Motif



Legend
Y = U or C
R = A or G

Figure 6: Change in Serum HBV DNA Levels Following 14 Days of Ribozyme Treatment of HBV Transgenic Mice

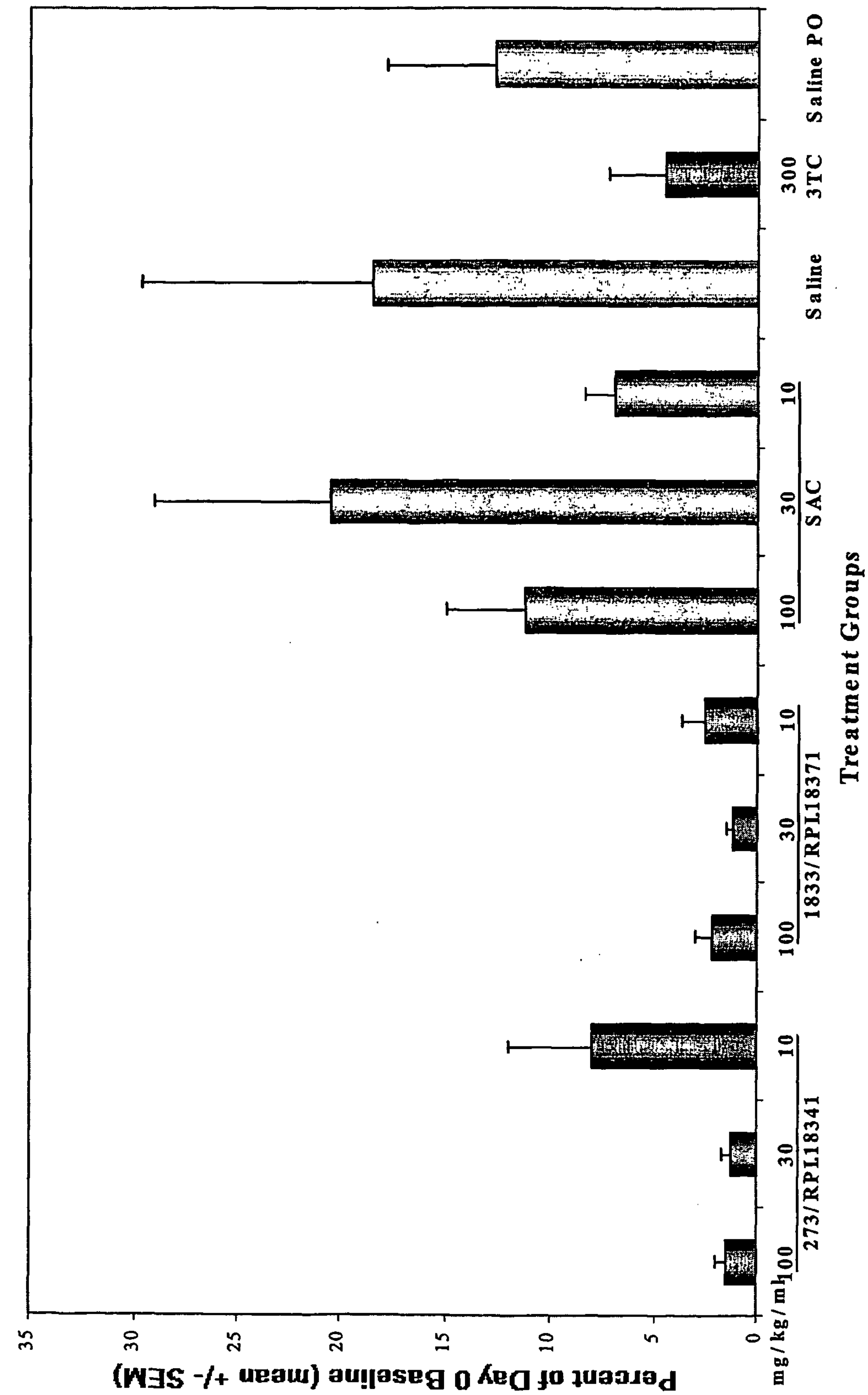


Figure 7: Mean Serum HBV DNA Levels Following 14 Days of Ribozyme Treatment of HBV Transgenic Mice

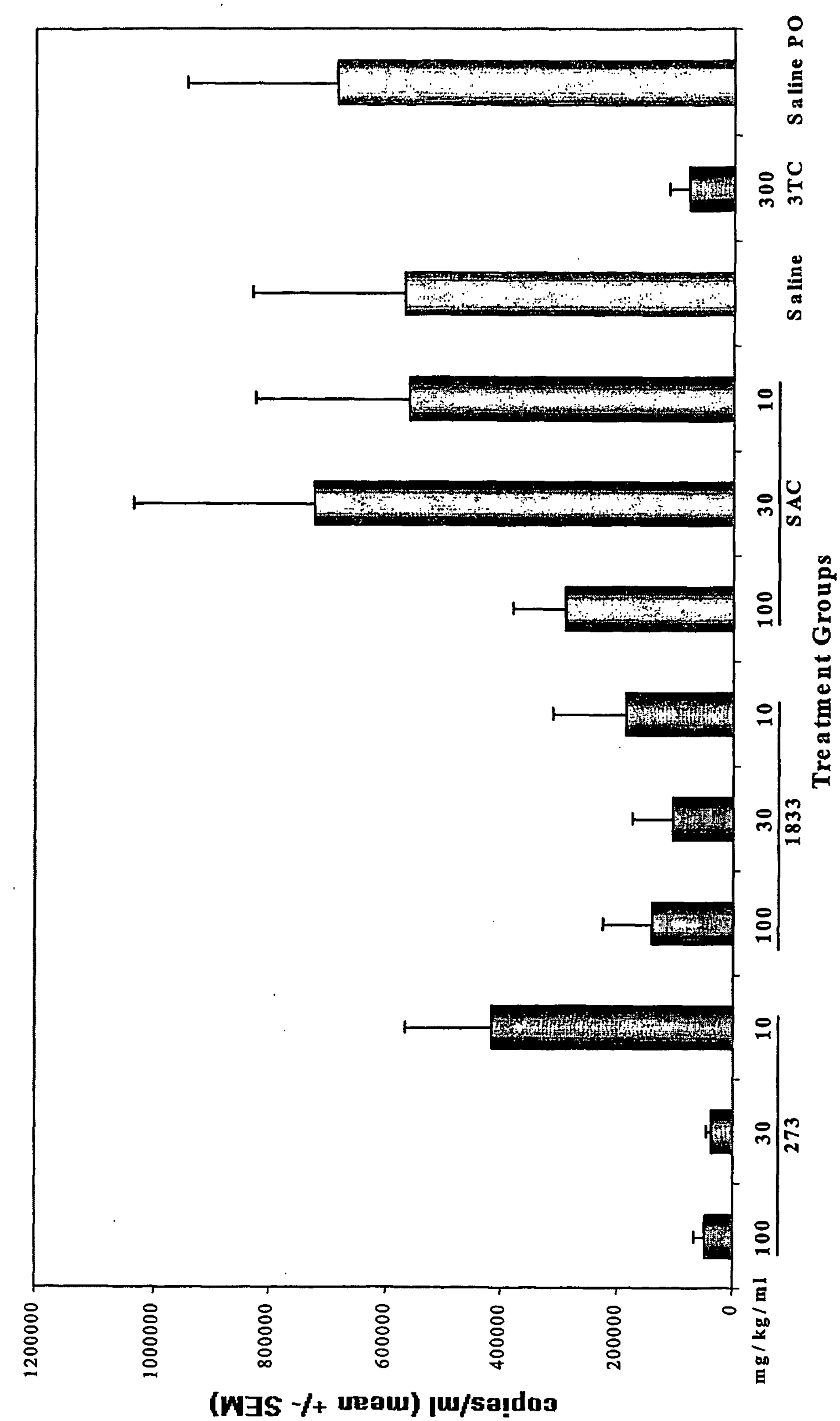


Figure 8: Change in Serum HBV DNA Levels (Log) Following 14 Days of Ribozyme Treatment of HBV Transgenic Mice

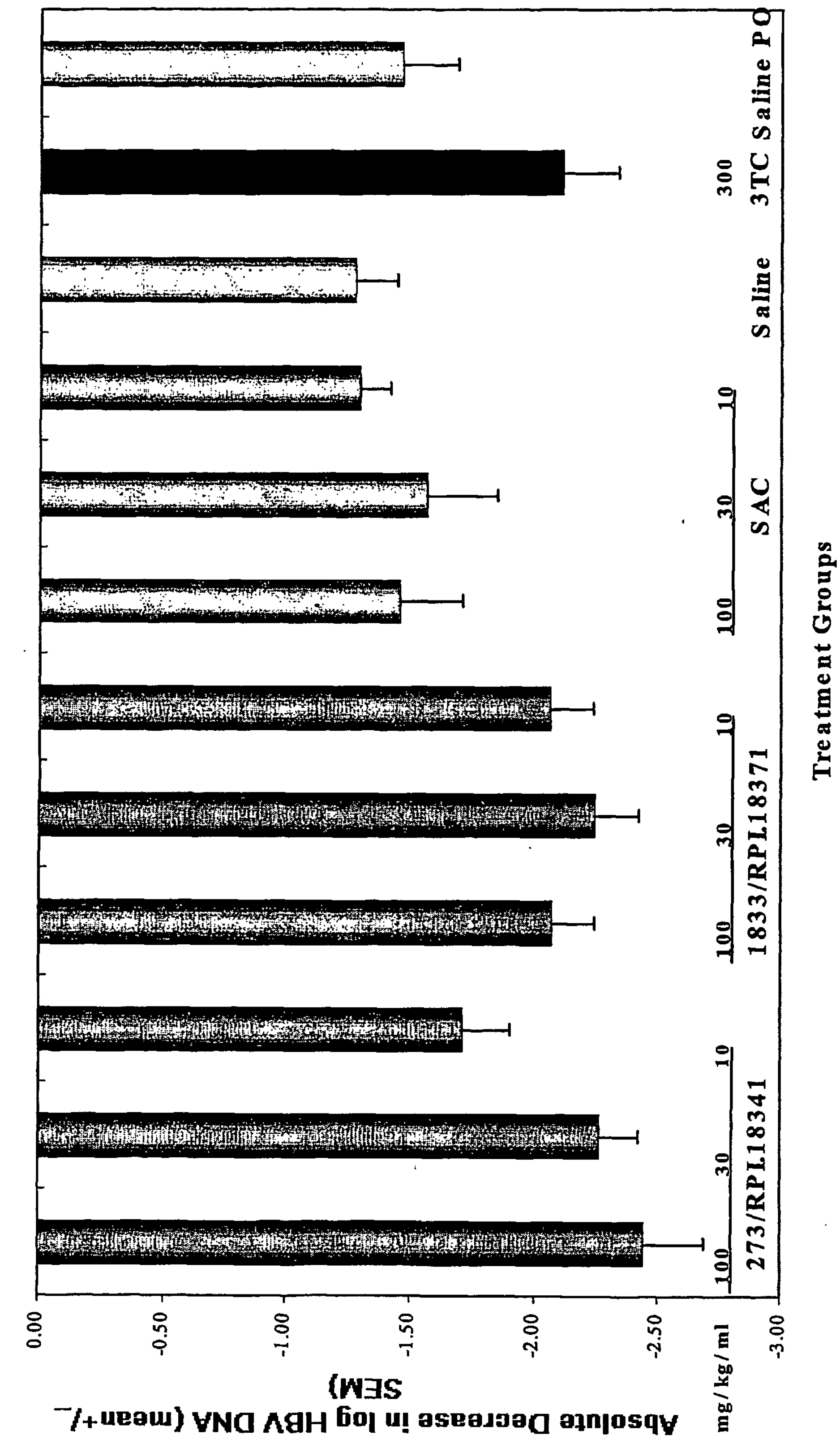


Figure 9: anti-HBV Ribozymes in HepG2.2.15 Cells: HBV DNA

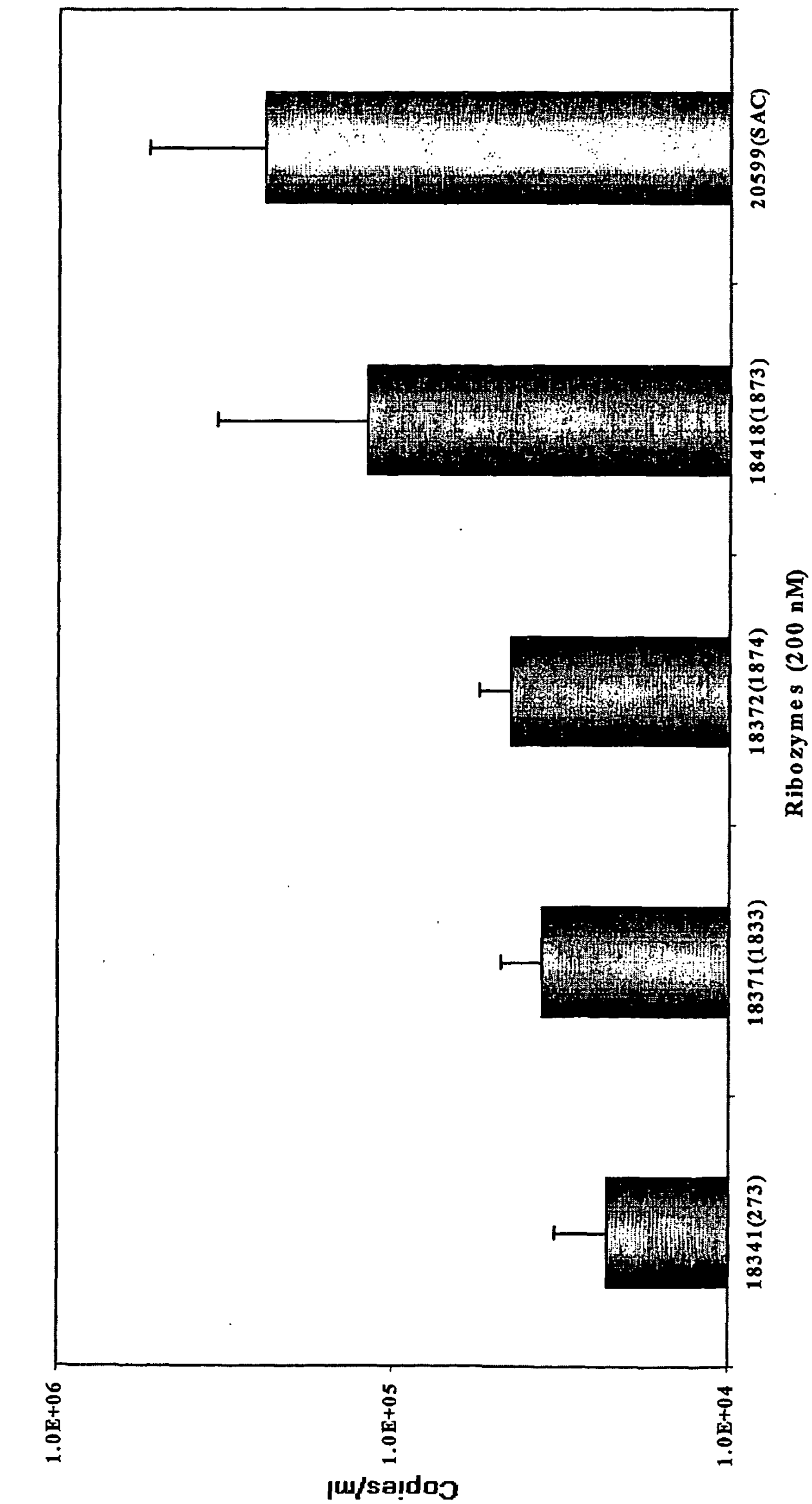


Figure 10: Arm, Loop, and Stem Variants of Anti-HBV Ribozyme Targeting Site 273: HBsAg Levels in Hep G2 Cells

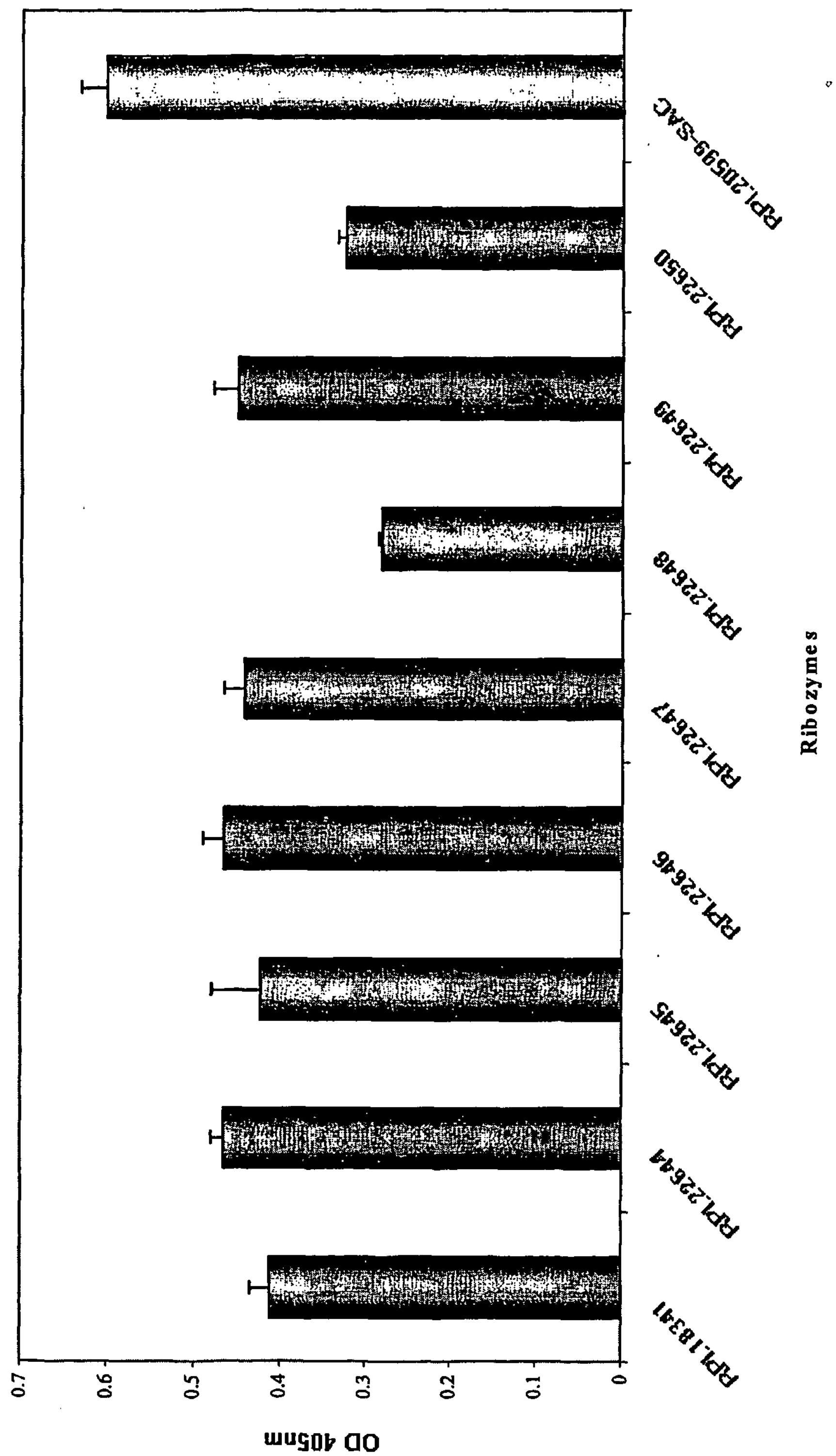


Fig 11: Hep G2 Cells Treated with RPI.18341 and Interferon: HBsAg ELISA

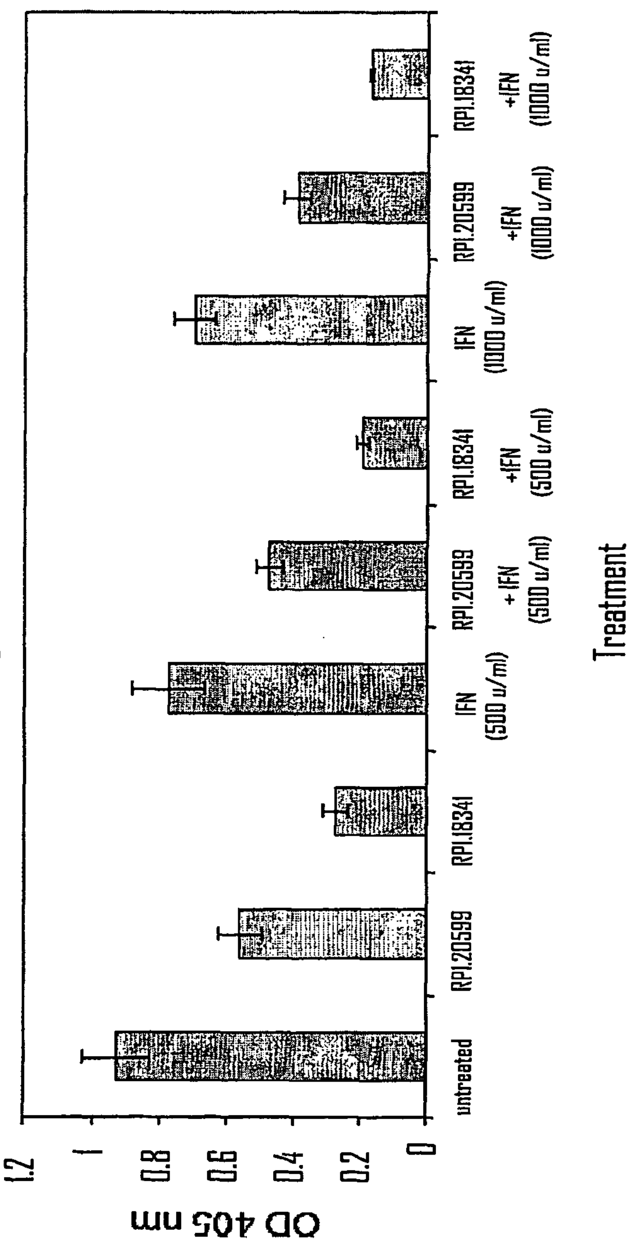


Fig 12: Hep G2 Cells Treated with 100 nM RPI.18341 and Lamivudine (3TC): HBsAg ELISA

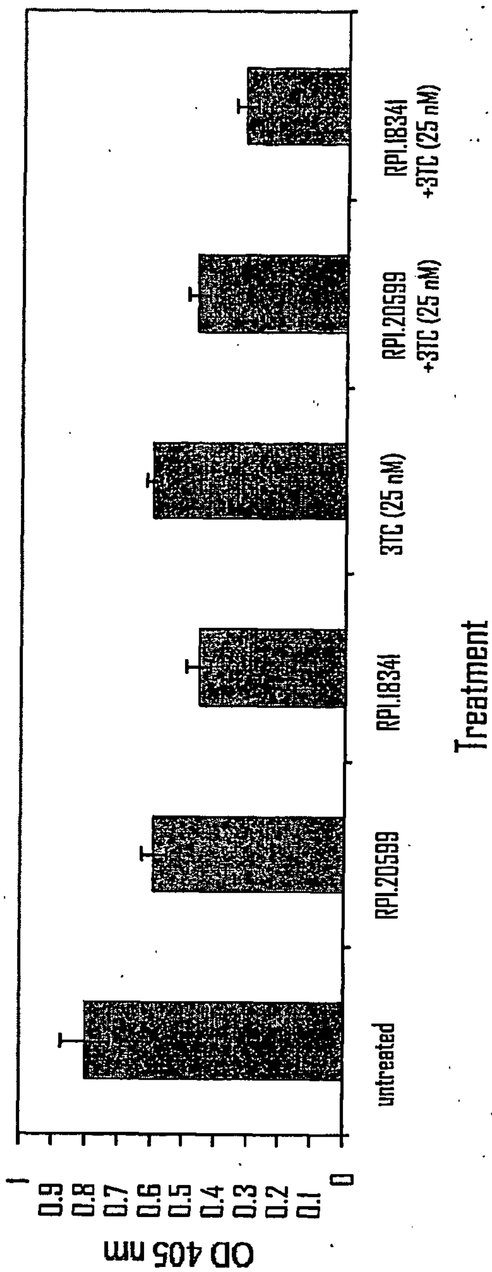


Figure 13: HBV Reverse Transcription

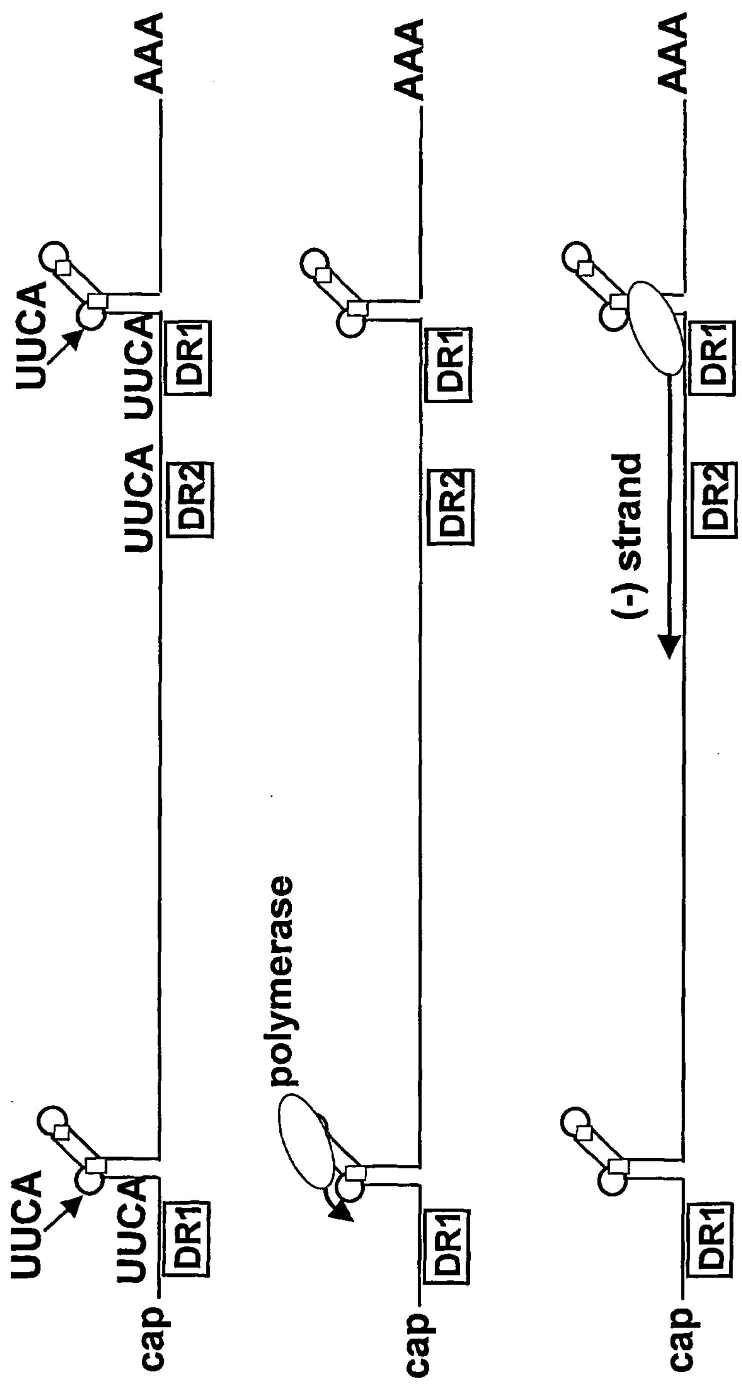


Figure 14: HBV RT Inhibition

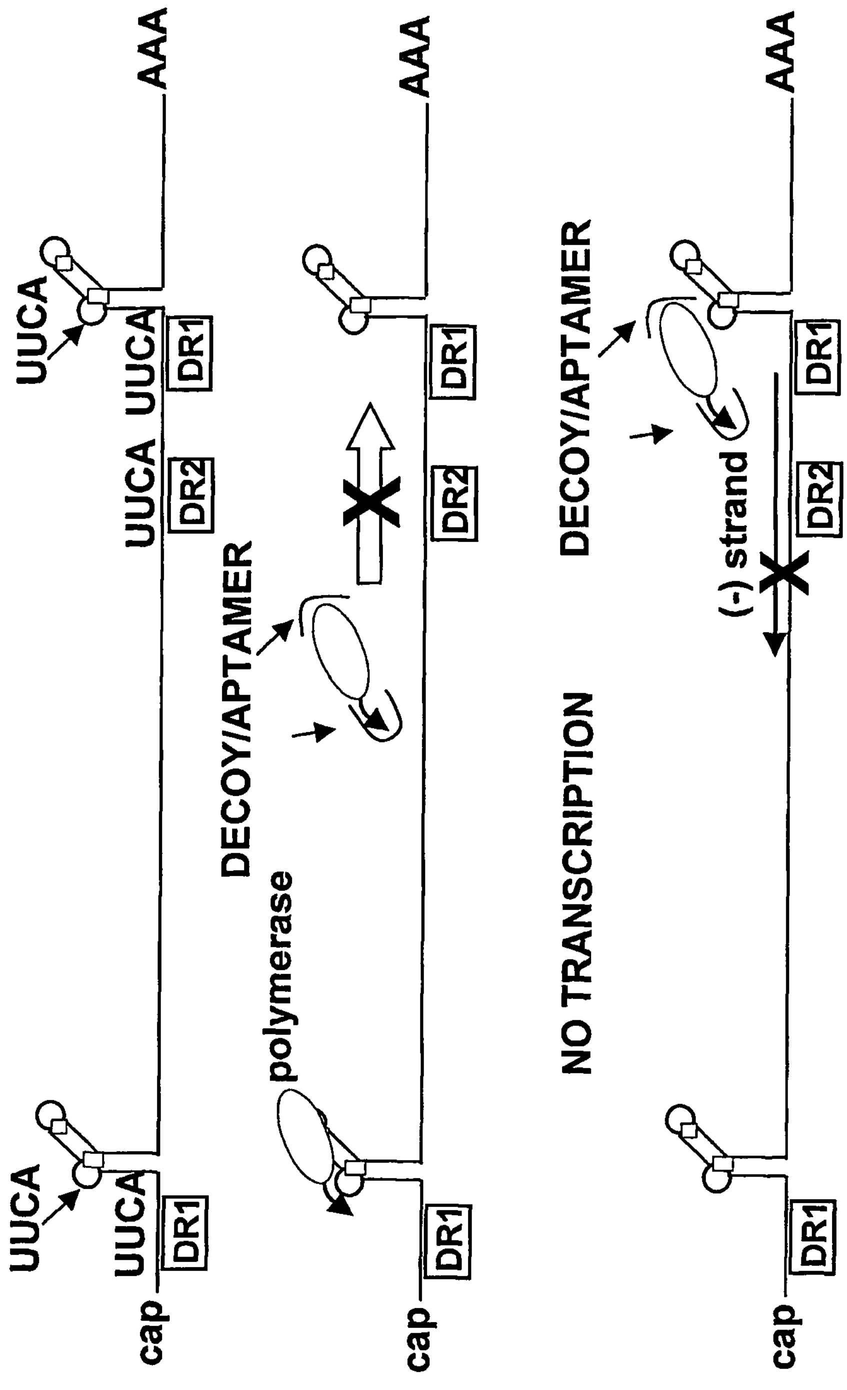


Figure 15: Screening of HBV RT Primer Competitive Inhibitors (2'-O-Allyl): HBsAg

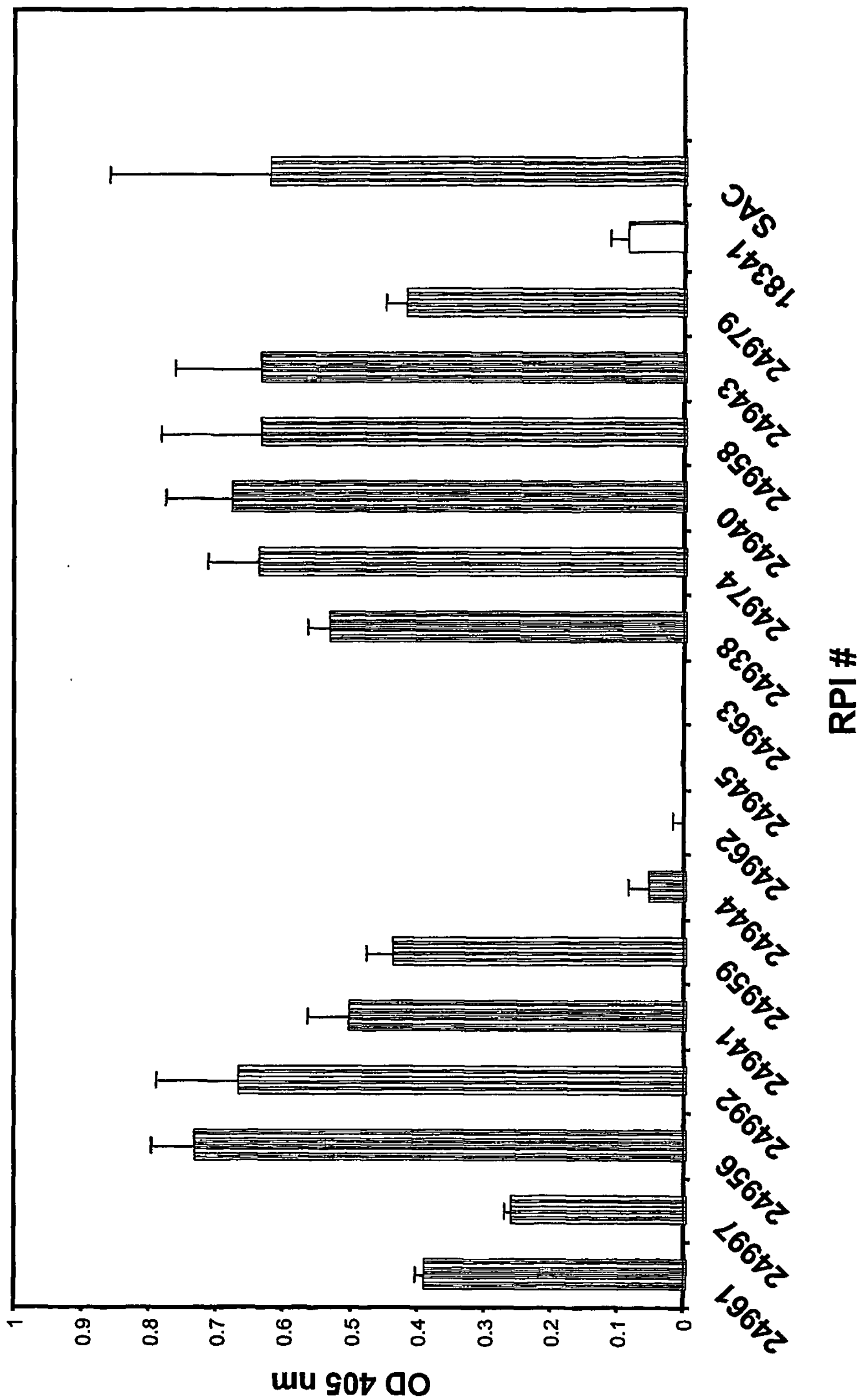


Figure 16: Screening of HBV RT Primer Competitive Inhibitors (2'-O-Methyl): HBsAg

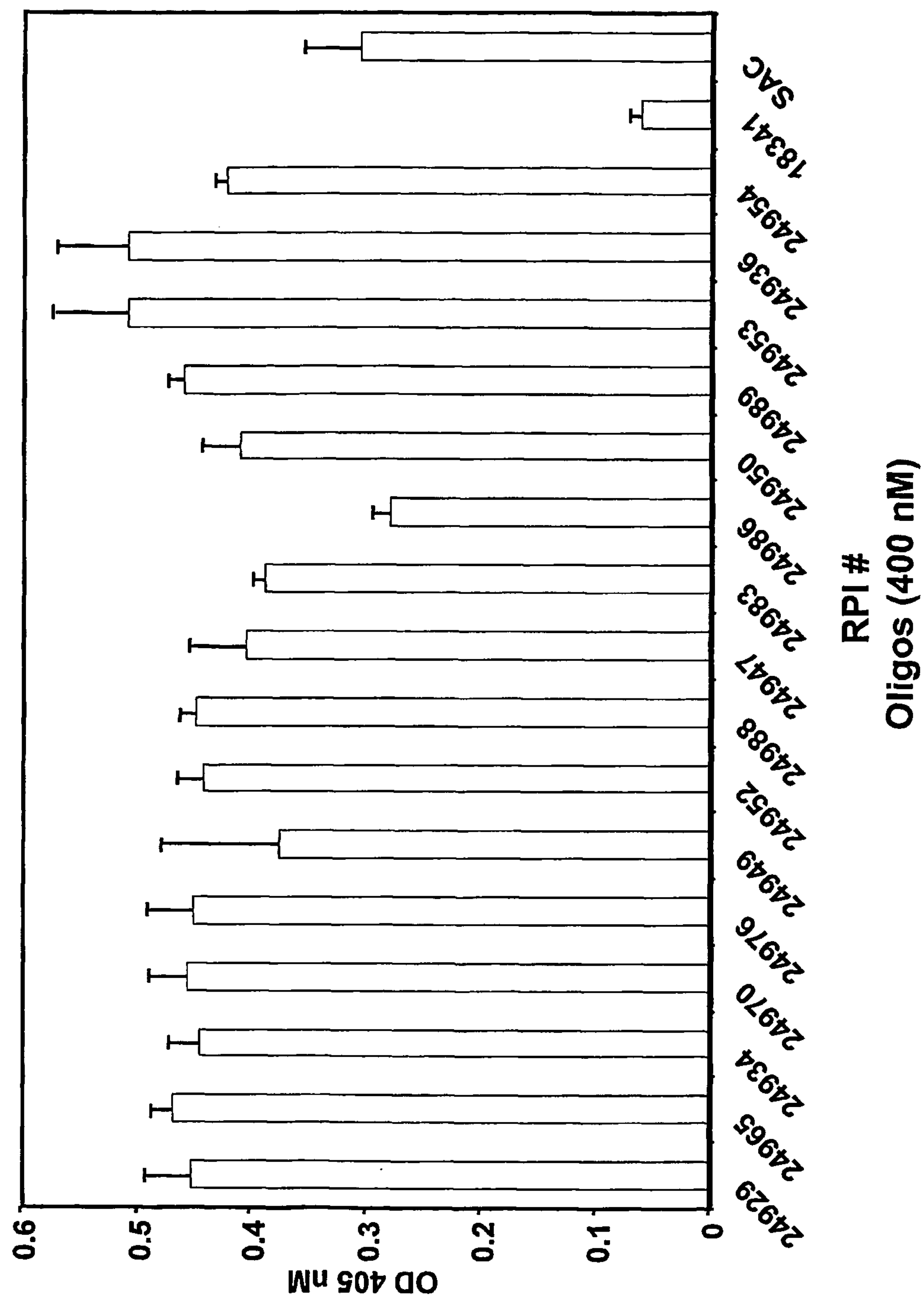


Figure 17: Dose Response with 2'-O-Methyl UUCAUUCA Oligo: HBsAg

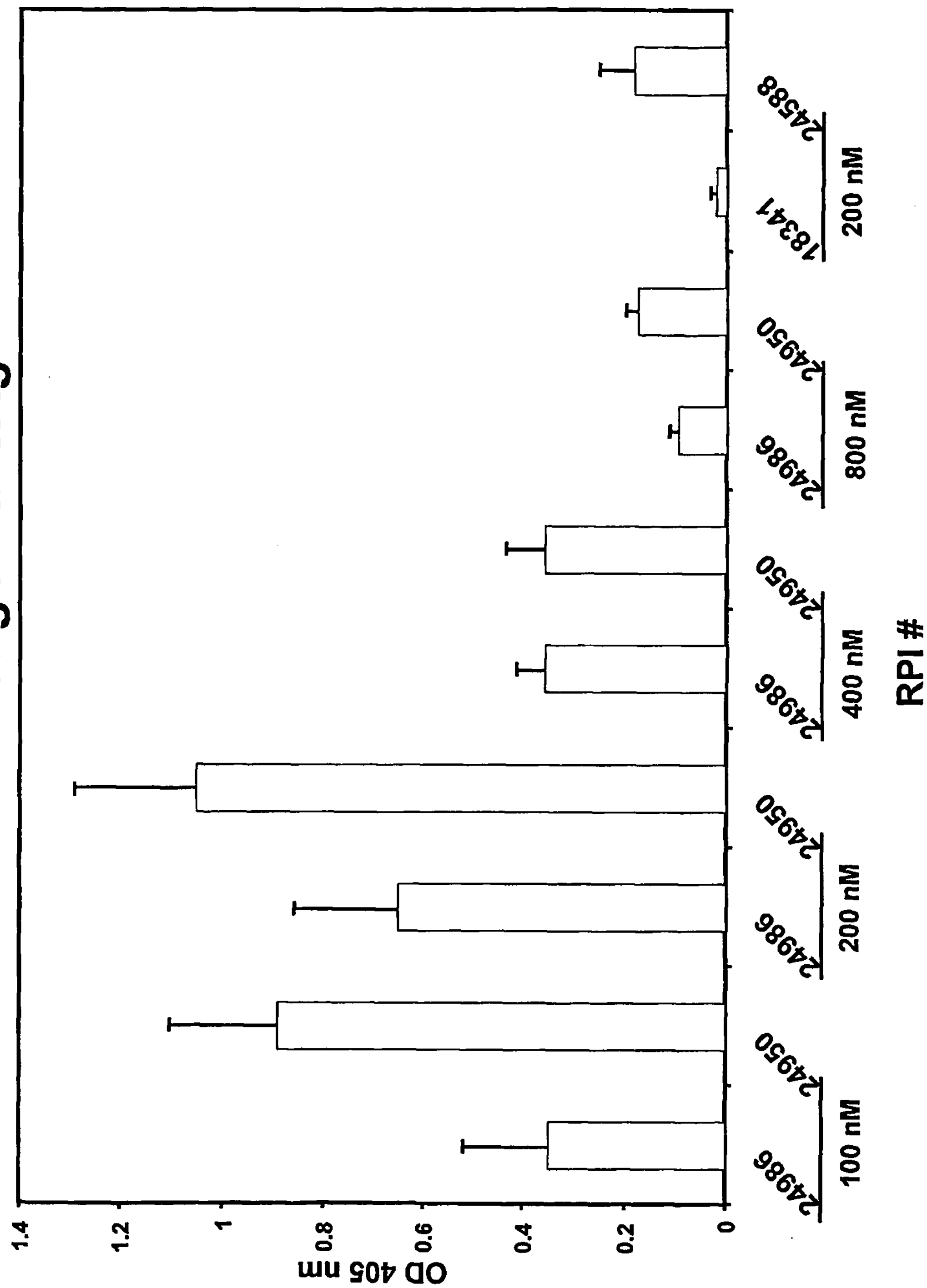


Figure 18: HBV Enhancer I Oligo Screen 200 nM:HBsAg

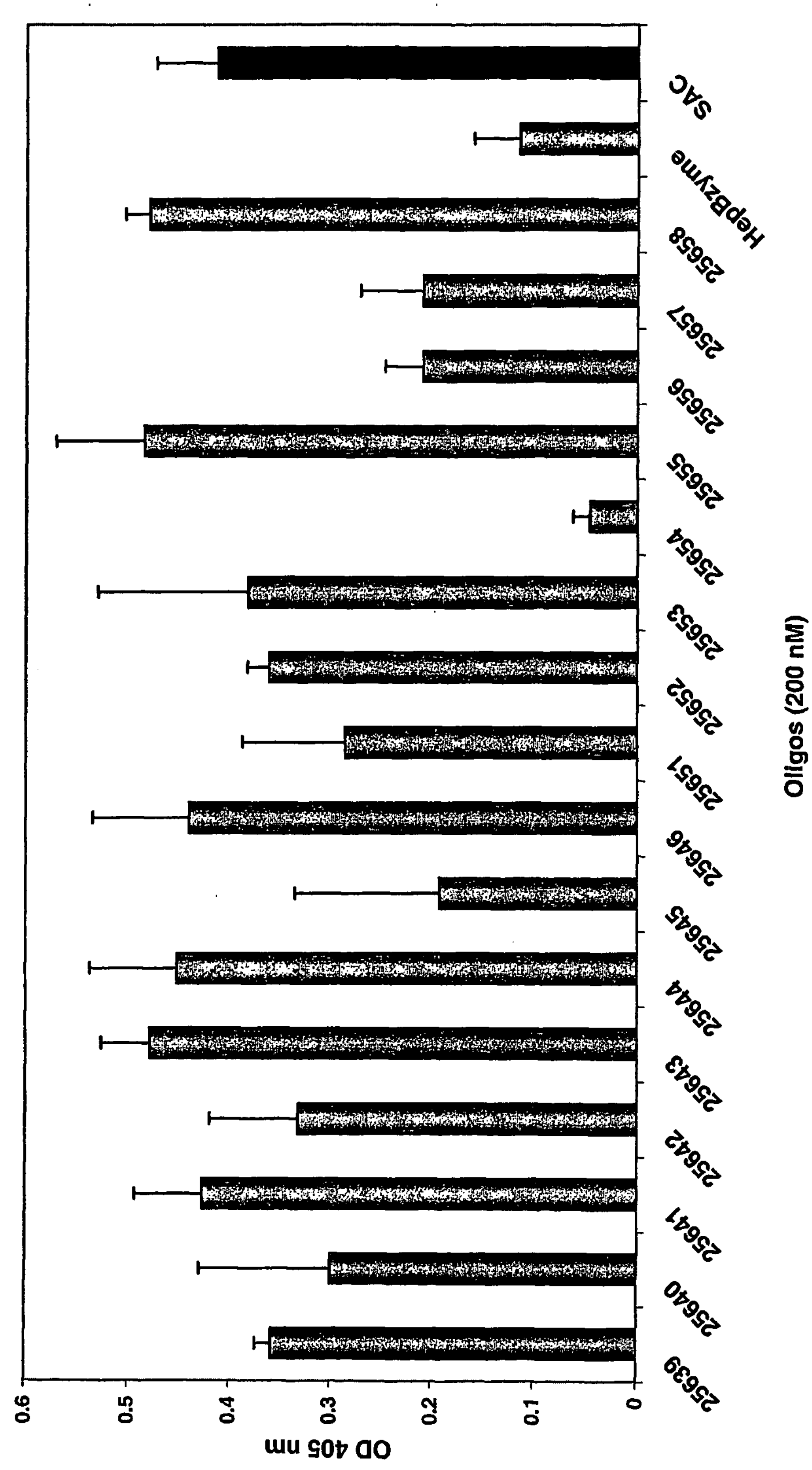


Figure 19: HBV Enhancer I Oligo Screen 400 nM: HBsAg

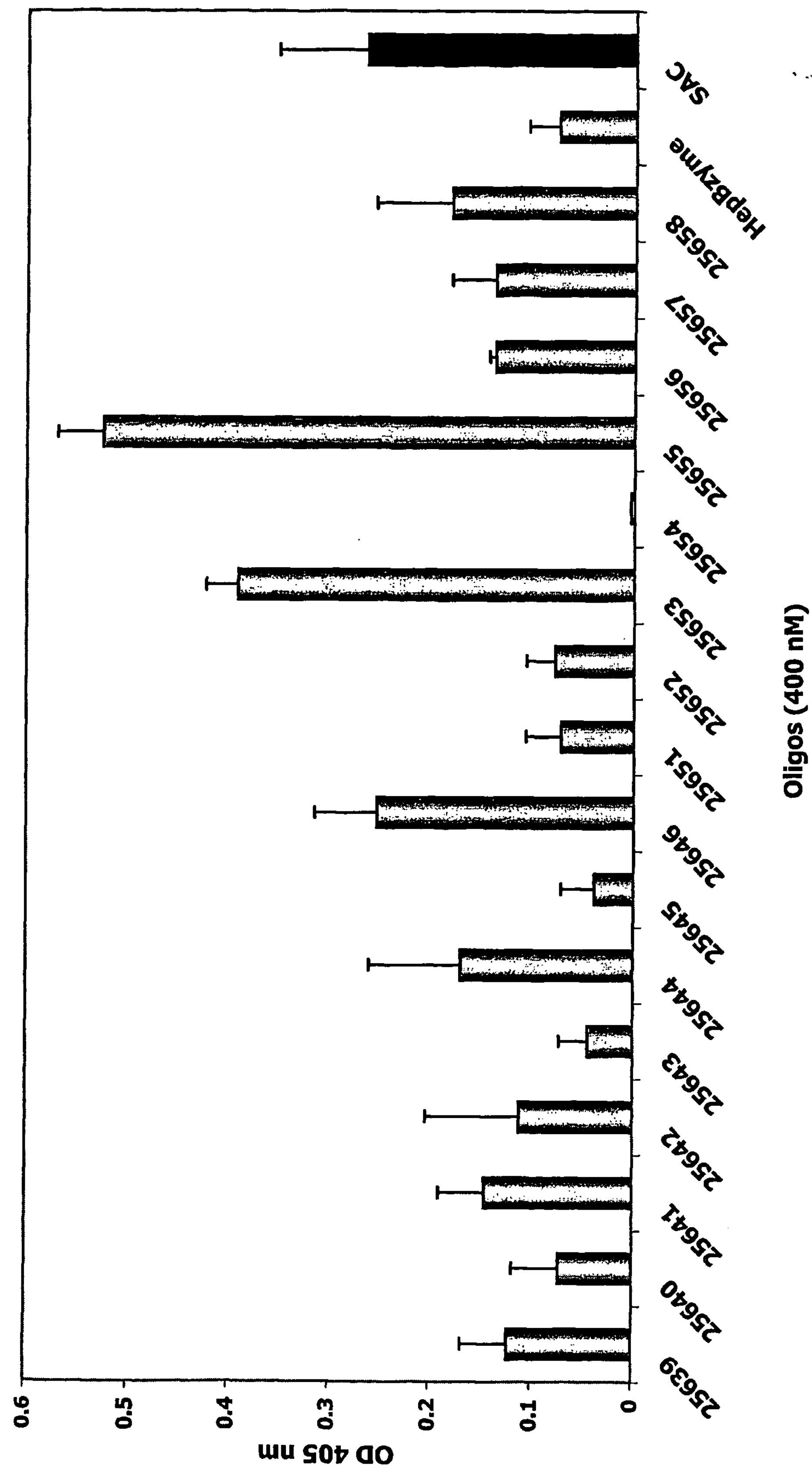


Figure 20: HBV Enhancer 1 Oligos Dose Response HBsAg

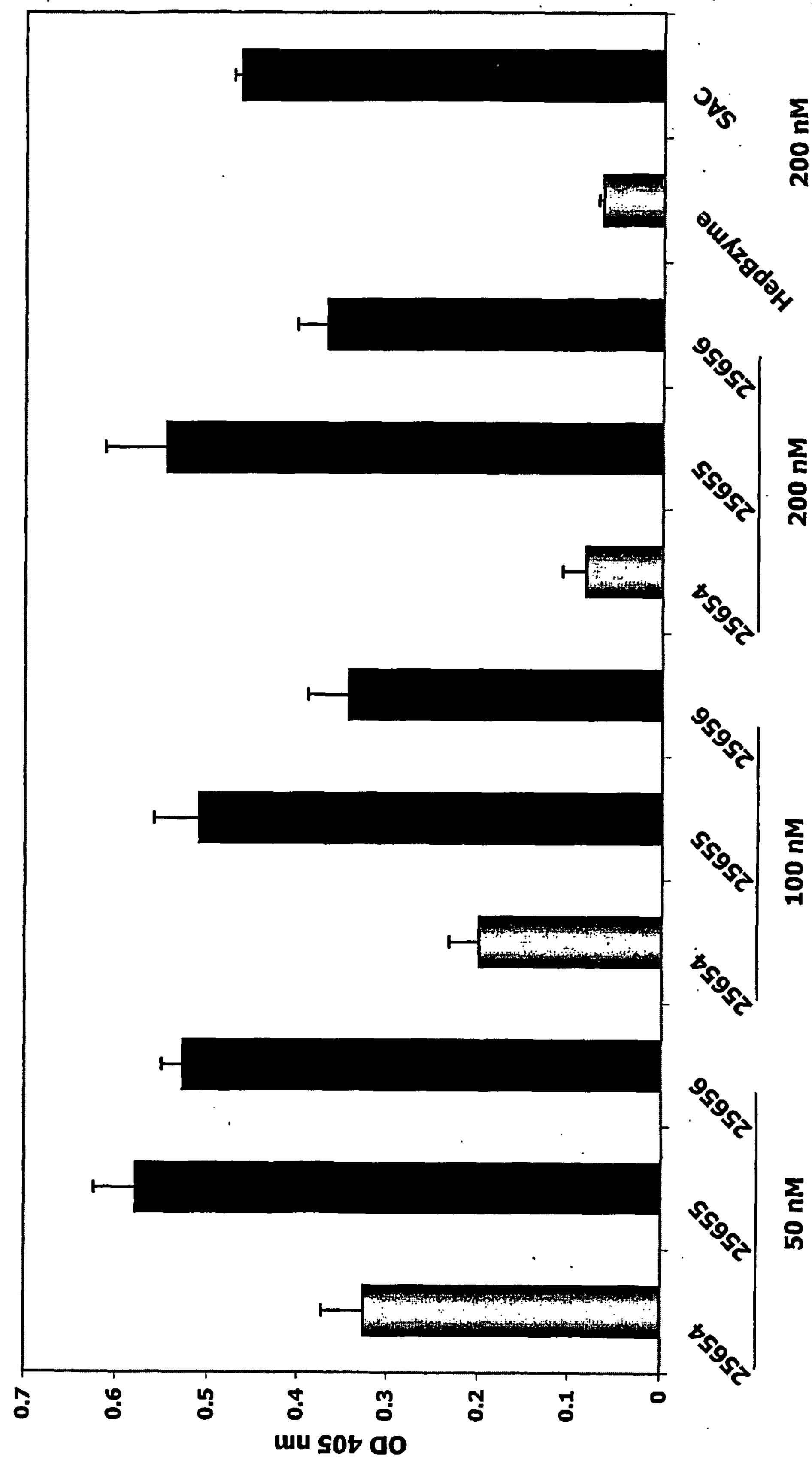


Figure 21: Growth of HepG2.2.15 tumors in Athymic Nu/Nu female mice

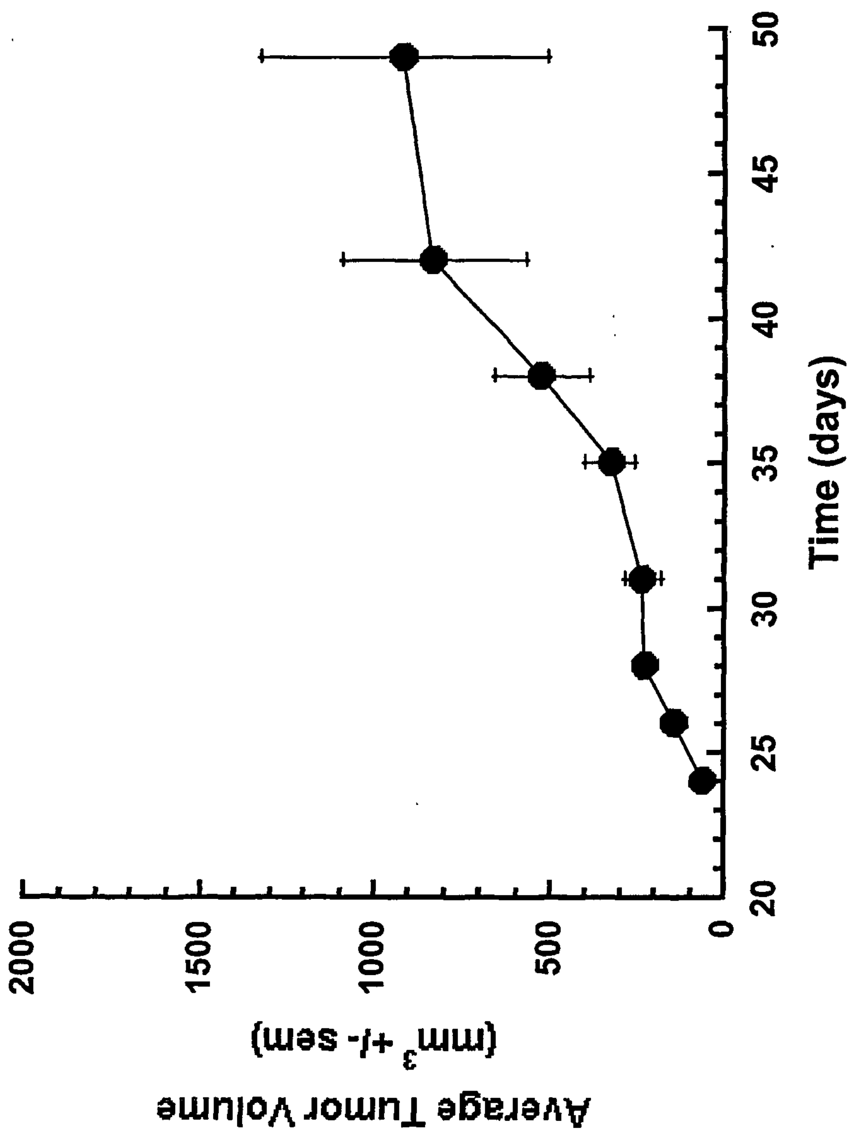


Figure 22: Growth of HepG2.2.15 tumors in Athymic Nu/Nu female mice

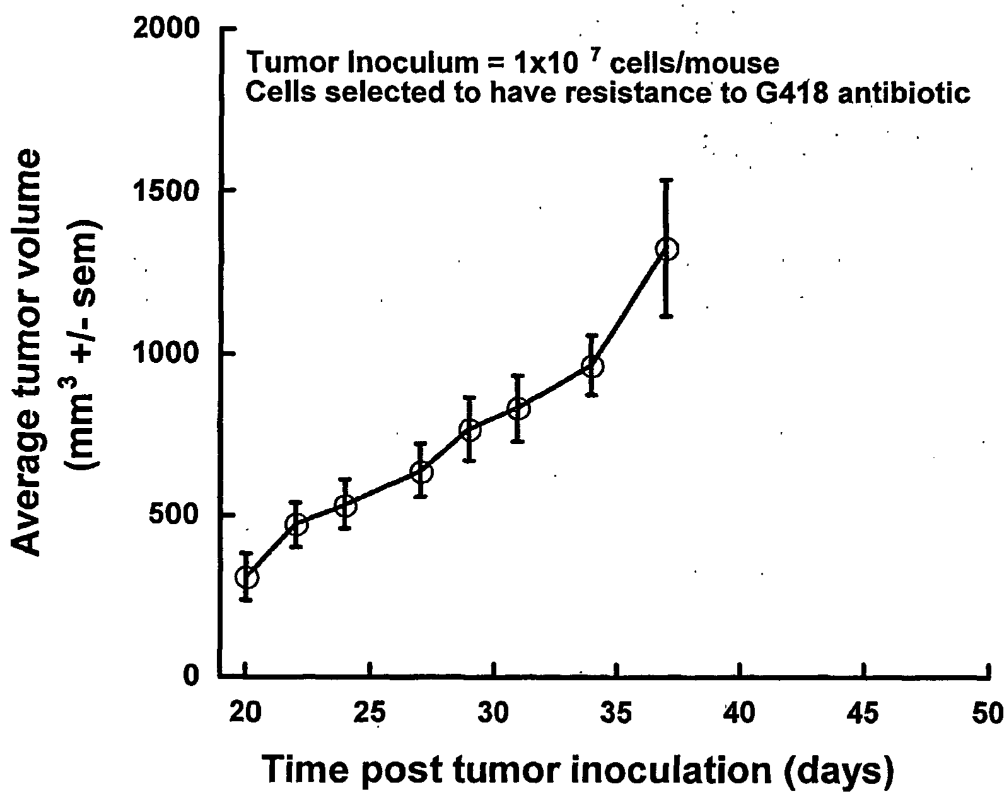


FIGURE 23 Dual Reporter System for Cytoplasmic HCV Target

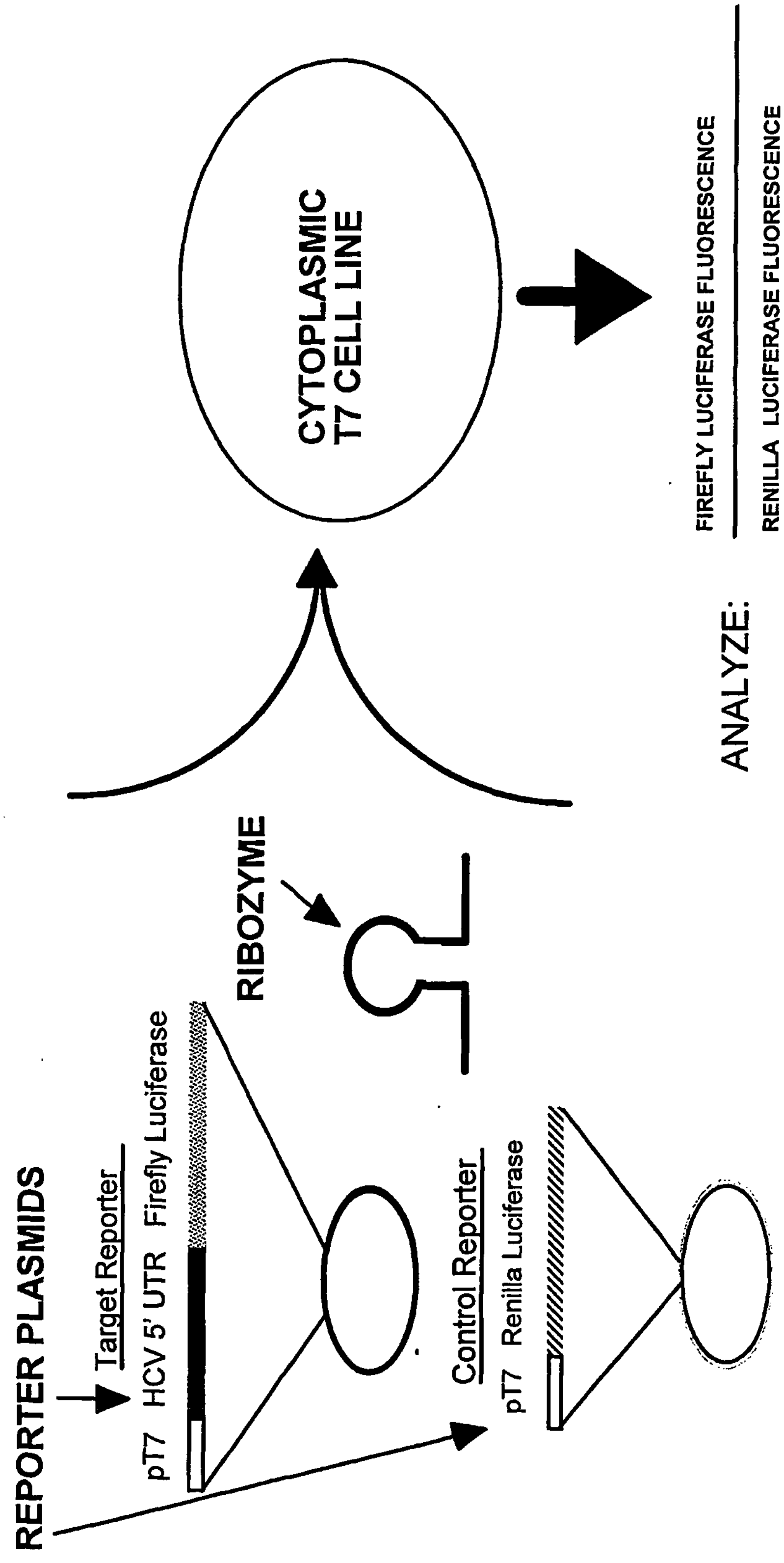


Figure 24: Secondary structure of the HCV 5'UTR

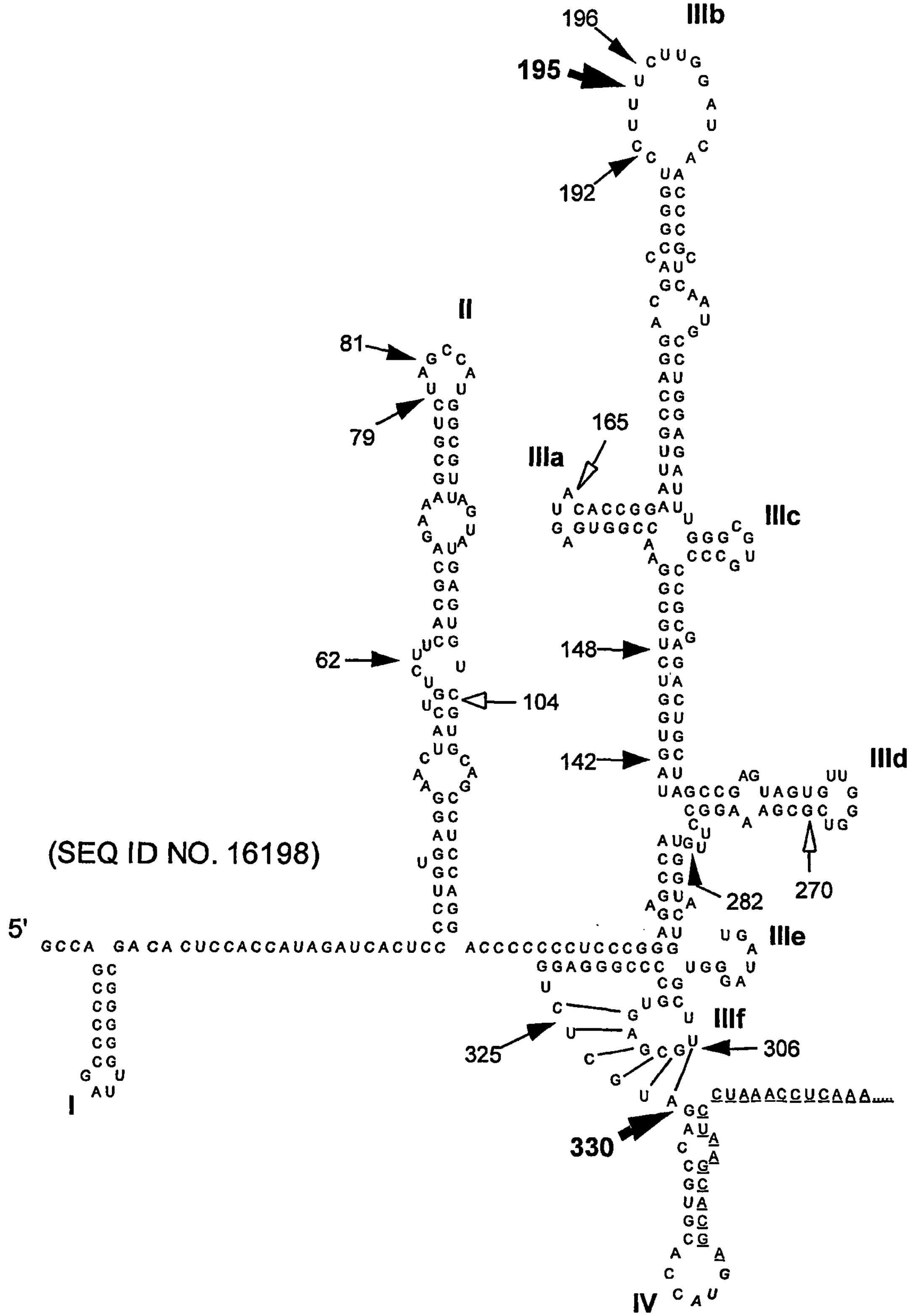
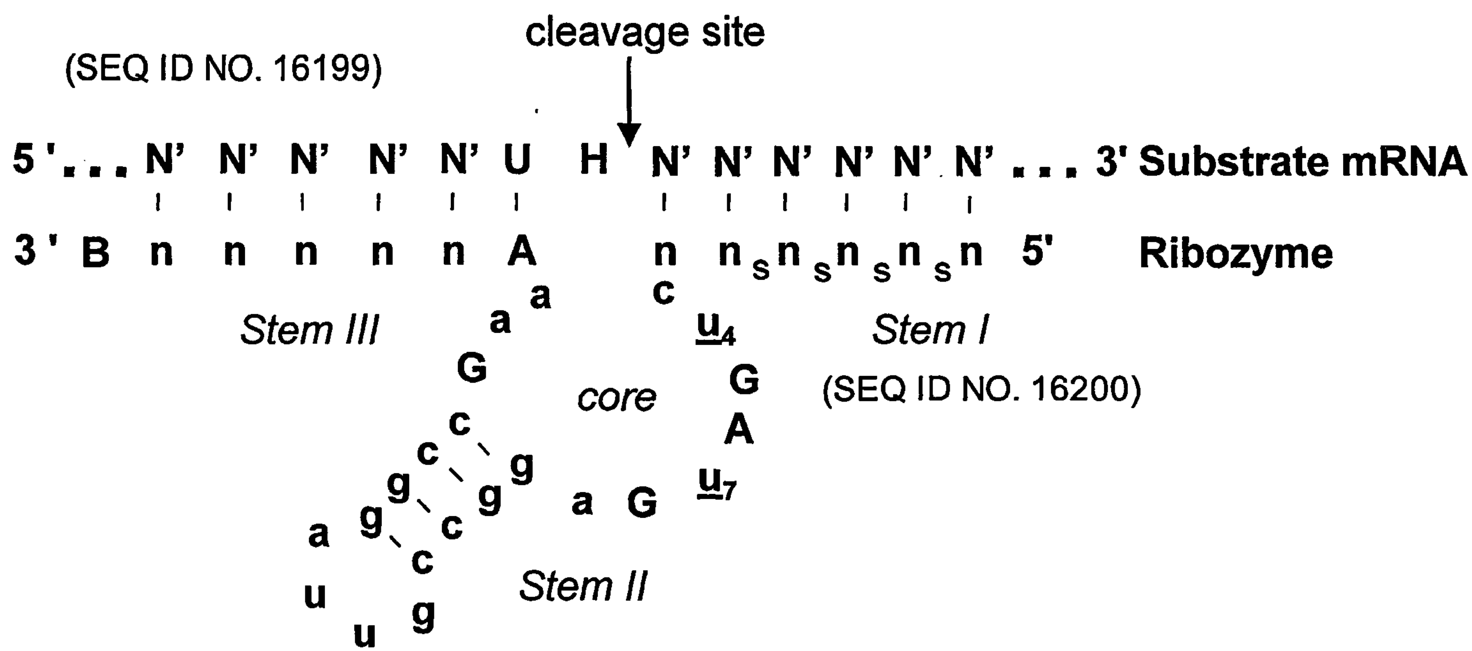


Figure 25: A Chemically Stabilized Enzymatic Nucleic Acid Molecule



UPPER CASE = RIBO nucleotide
lower case = 2'-O-methyl nucleitide
u = 2'-deoxy-2'-amino Uridine
s = phosphorothioate
B = inverted deoxyabasic moiety

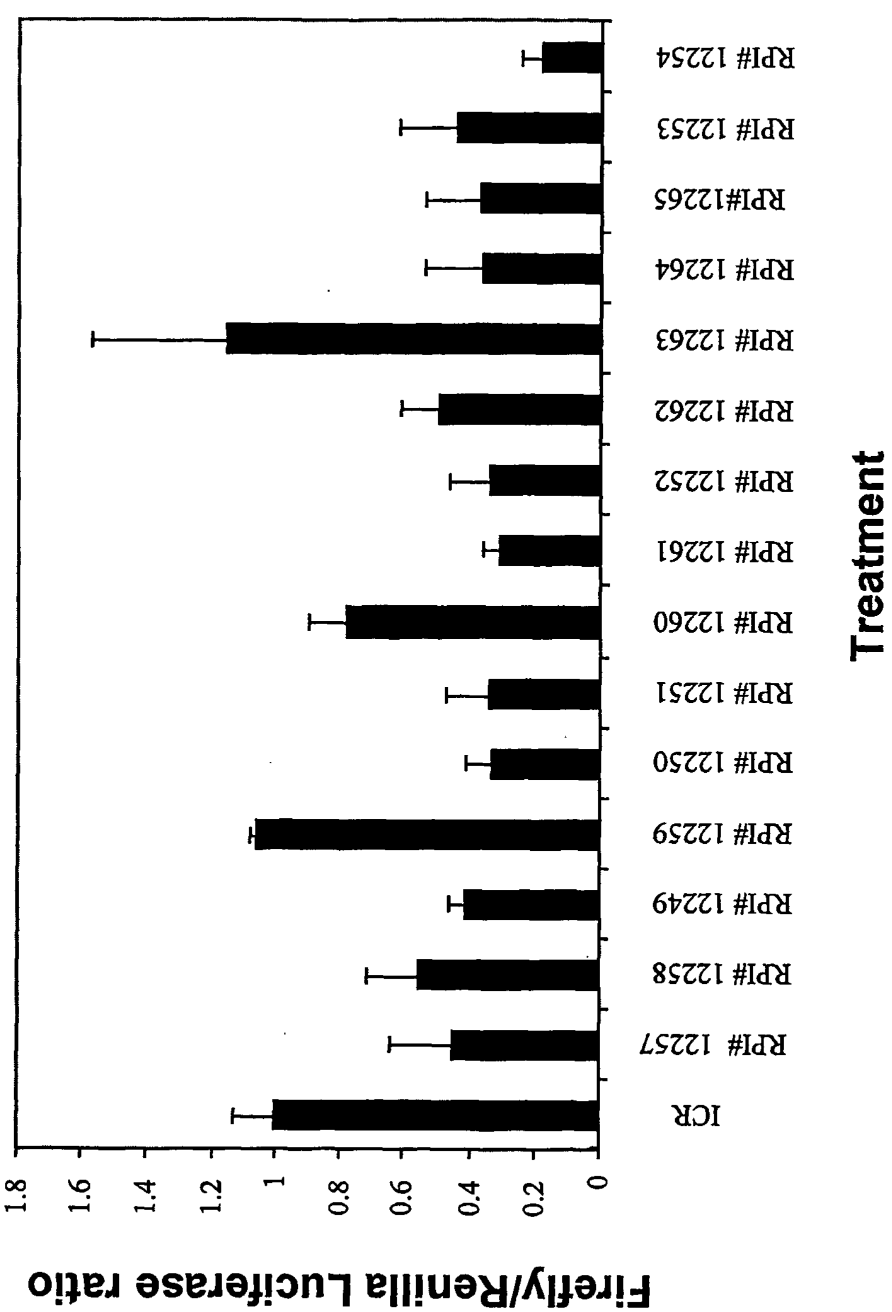


Figure 26A: Enzymatic nucleic acid mediated inhibition of HCV-luciferase expression

Figure 26B: Enzymatic nucleic acid mediated inhibition of HCV-luciferase expression

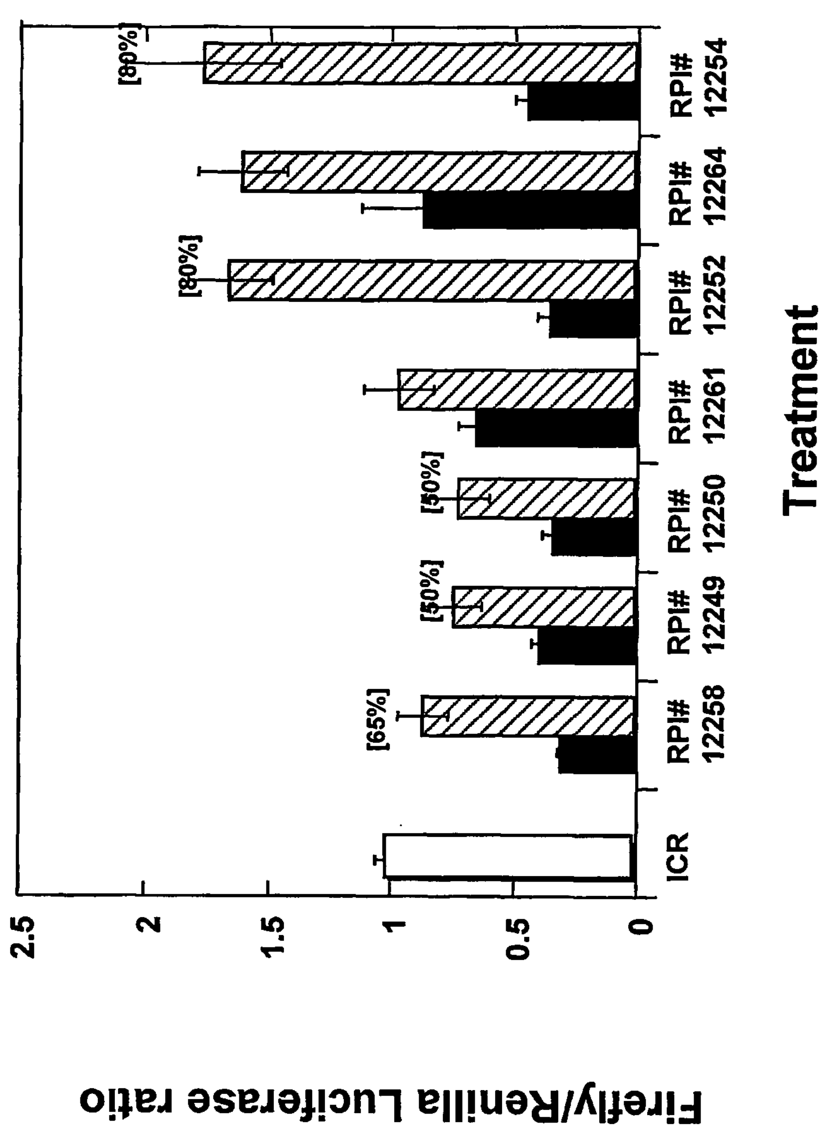


Figure 27A: Dose-dependent enzymatic nucleic acid inhibition of HCV/luciferase expression

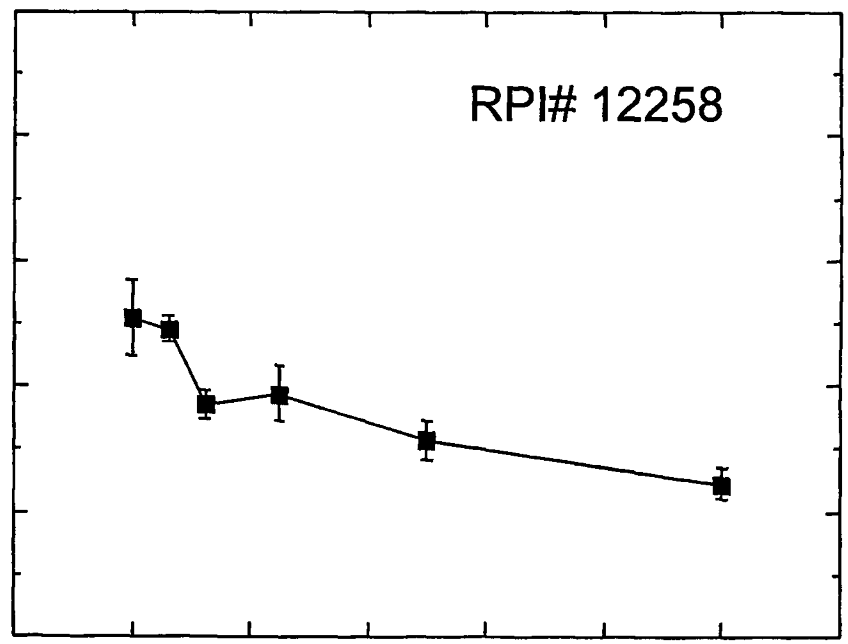


Figure 27B: Dose-dependent enzymatic nucleic acid inhibition of HCV/luciferase expression

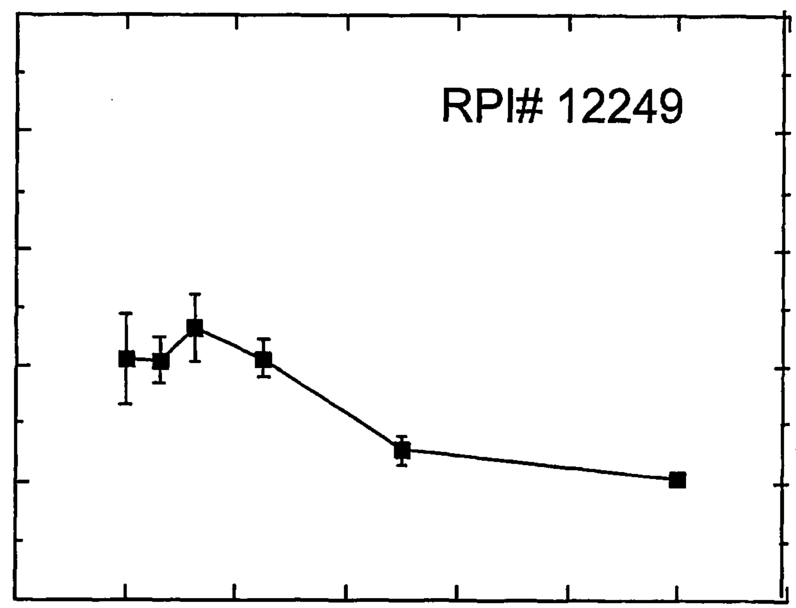


Figure 27C: Dose-dependent enzymatic nucleic acid inhibition of HCV/luciferase expression

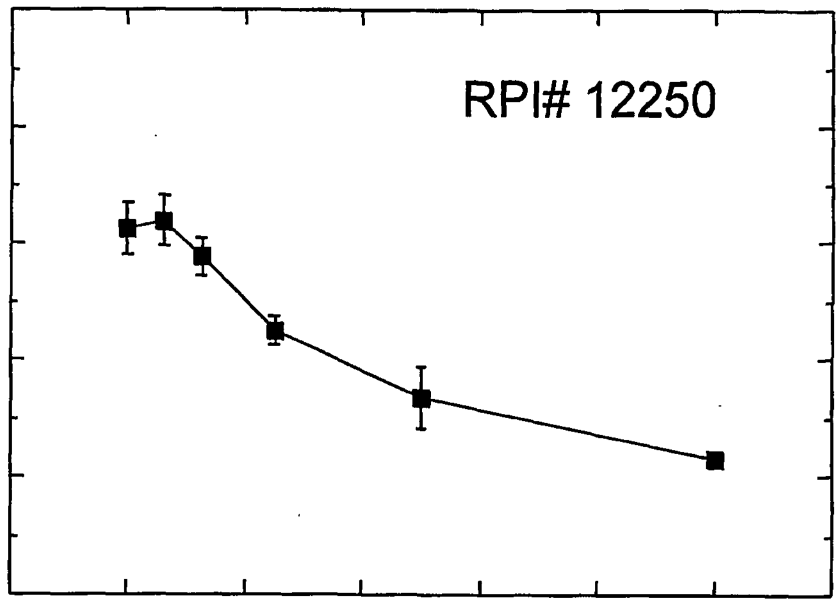


Figure 27D: Dose-dependent enzymatic nucleic acid inhibition of HCV/luciferase expression

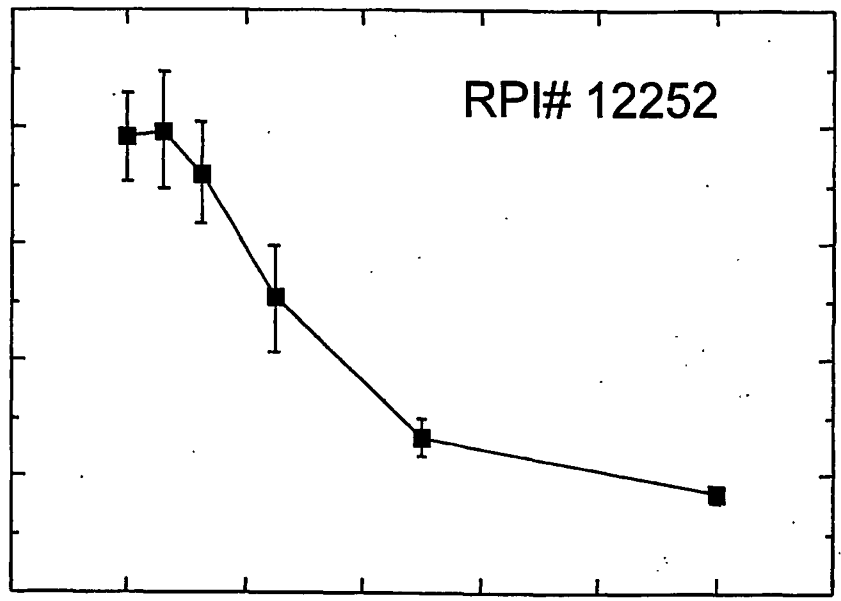
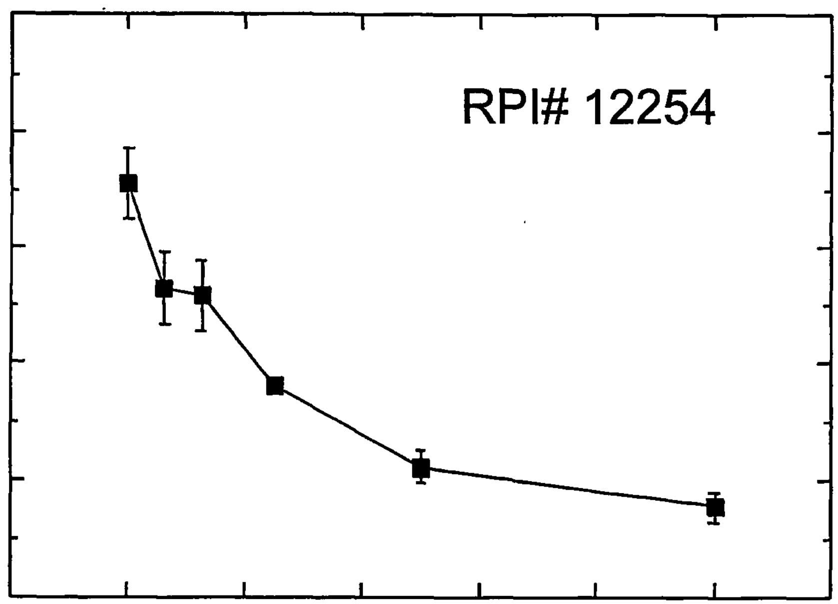


Figure 27E: Dose-dependent enzymatic nucleic acid inhibition of HCV/luciferase expression



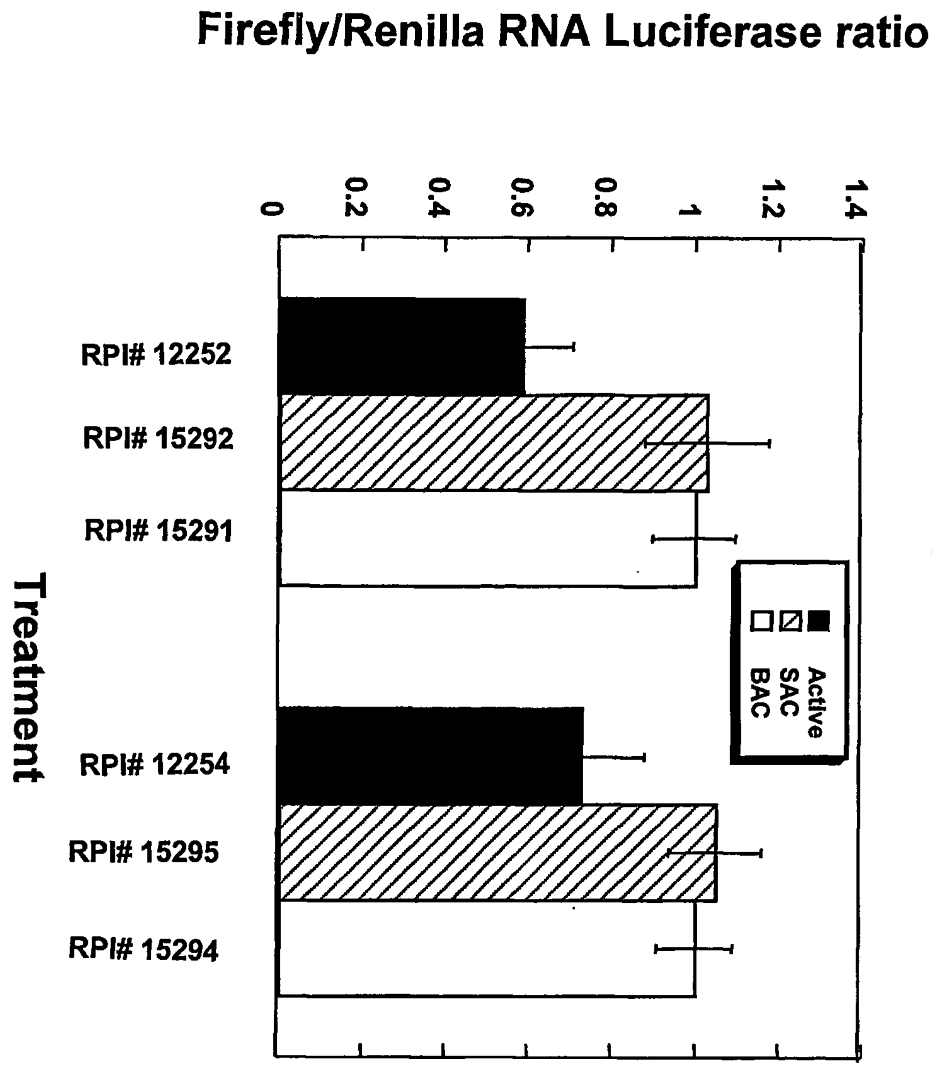


Figure 28A: Enzymatic nucleic acid reduction of HCV/luciferase RNA and inhibition of HCV-luciferase expression

Figure 28B: Enzymatic nucleic acid reduction of HCV/luciferase RNA and inhibition of HCV-luciferase expression

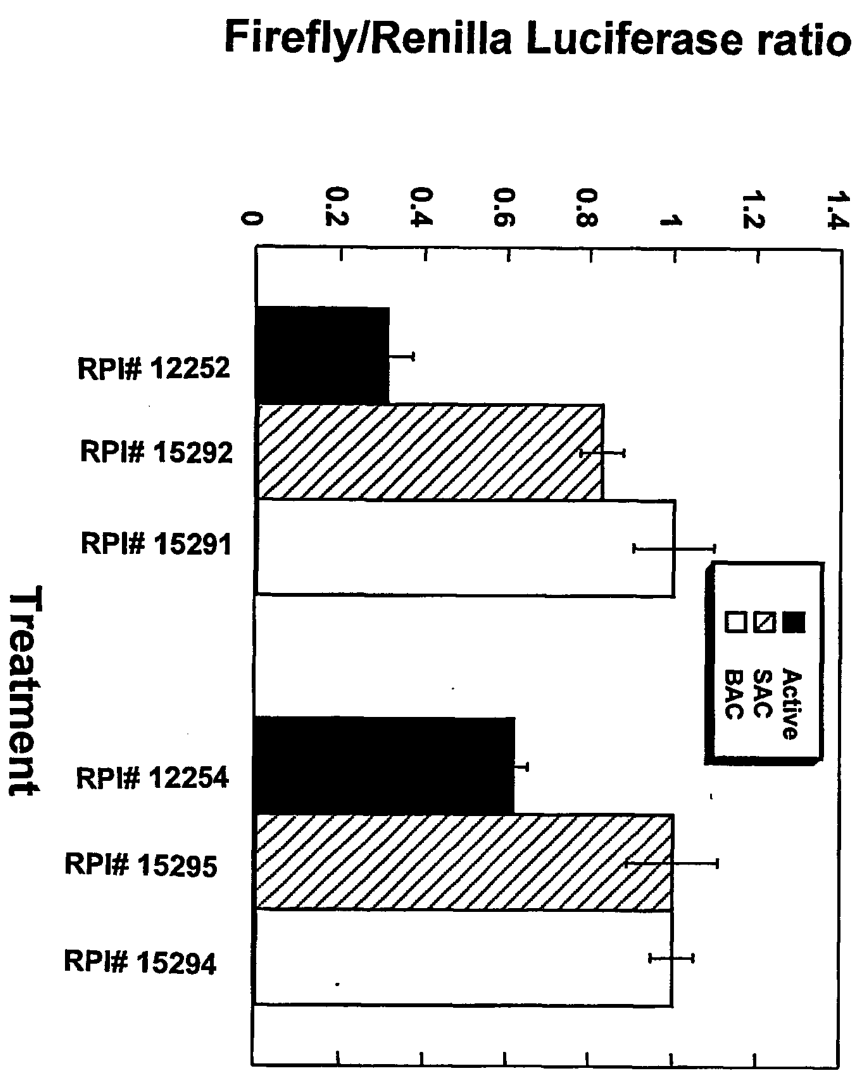


Figure 29A: Interferon Dose response with Enzymatic Nucleic Acid

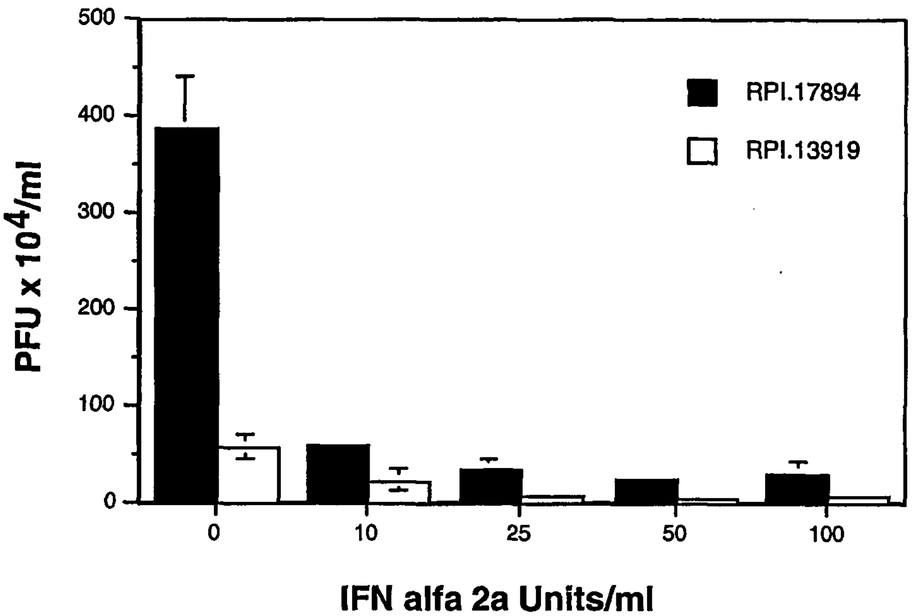


Figure 29B: Interferon Dose response with Enzymatic Nucleic Acid

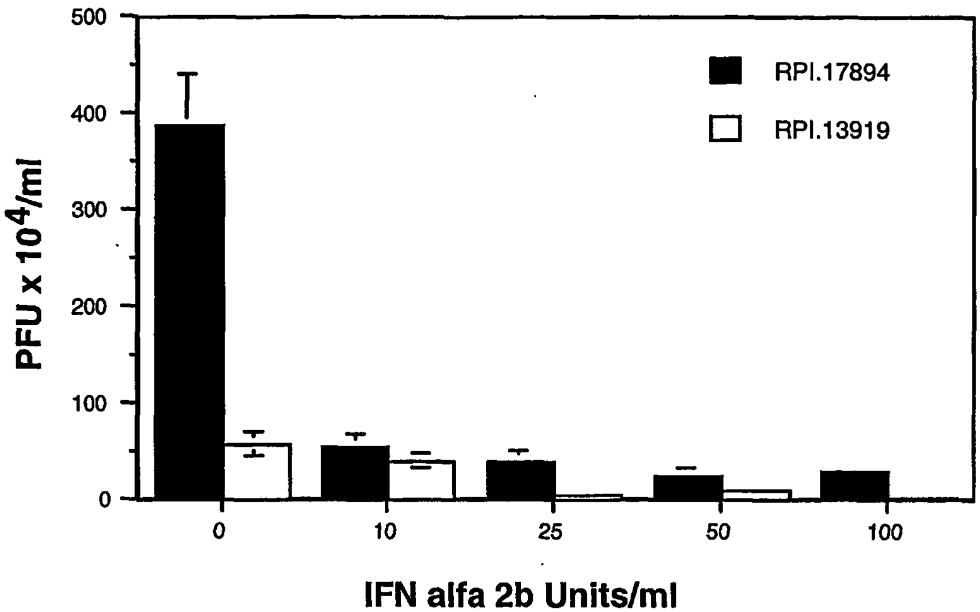


Figure 30: Site 195 anti-HCV enzymatic nucleic acid dose response in combination with interferon pretreatment

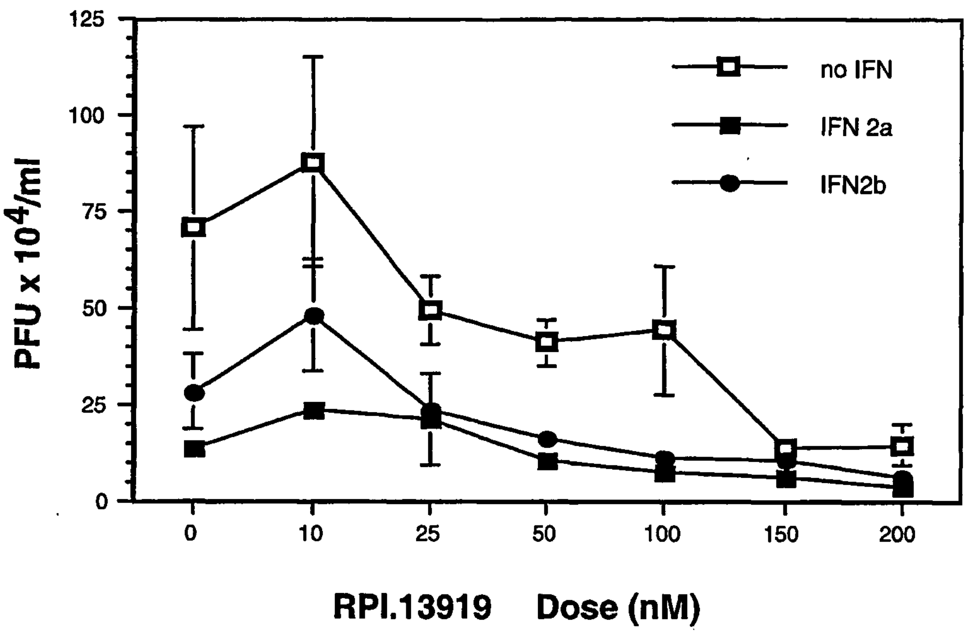


Figure 31A: CIFN dose response with site 195 anti-HCV enzymatic nucleic acid treatment

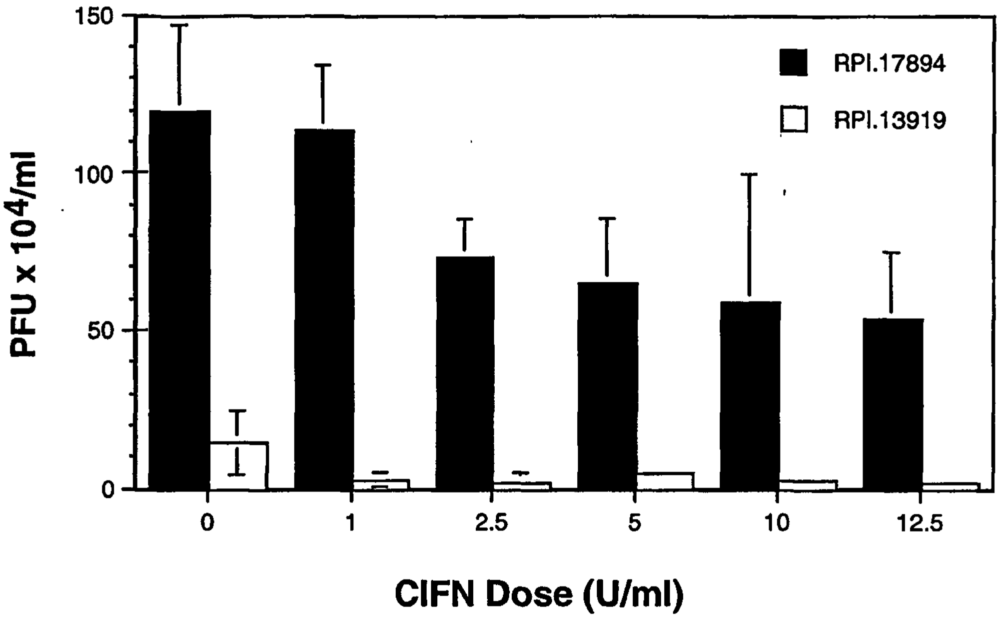


Figure 31B: Site 195 anti-HCV enzymatic nucleic acid dose response with CIFN pretreatment

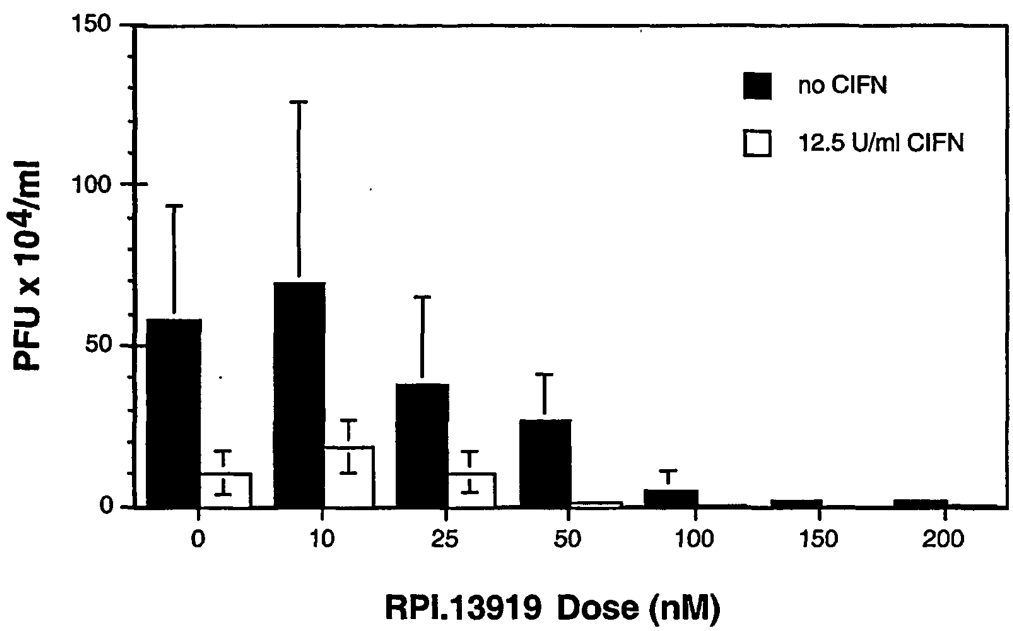


Figure 32: Enhanced antiviral effect of an anti-HCV enzymatic nucleic acid targeting site 195 used in combination with consensus interferon (CIFN)

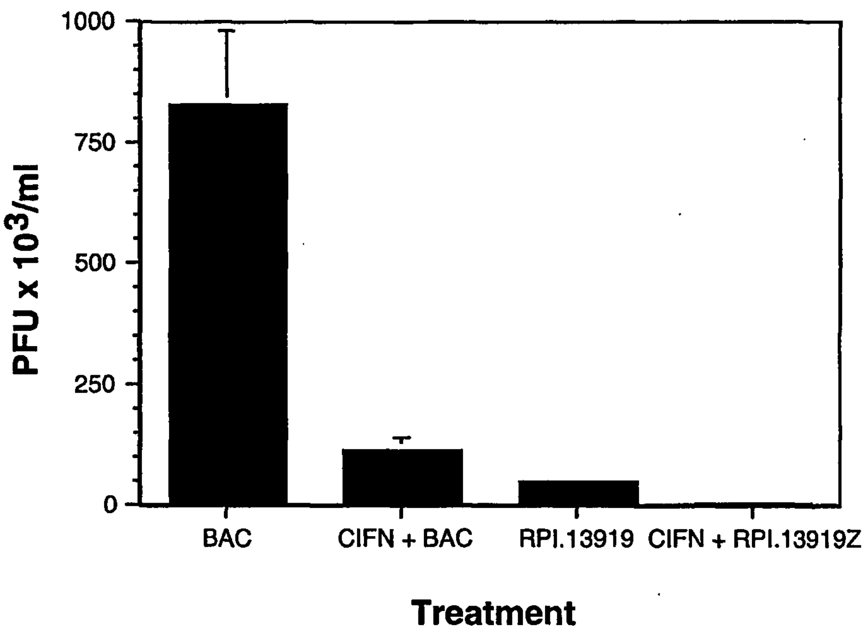
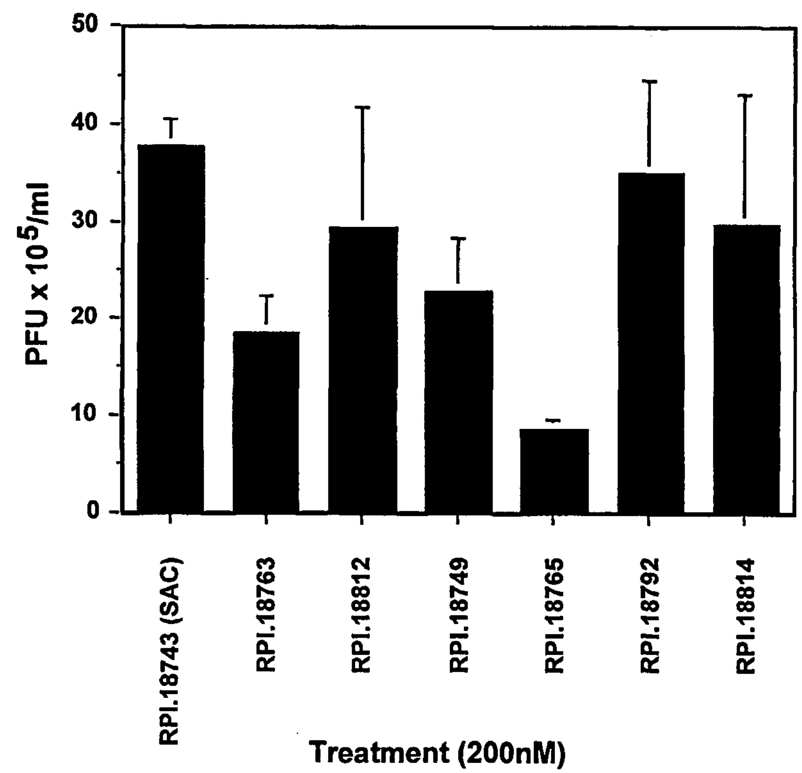


Figure 33: Inhibition of HCV-PV Replication by Zinzyme Treatment



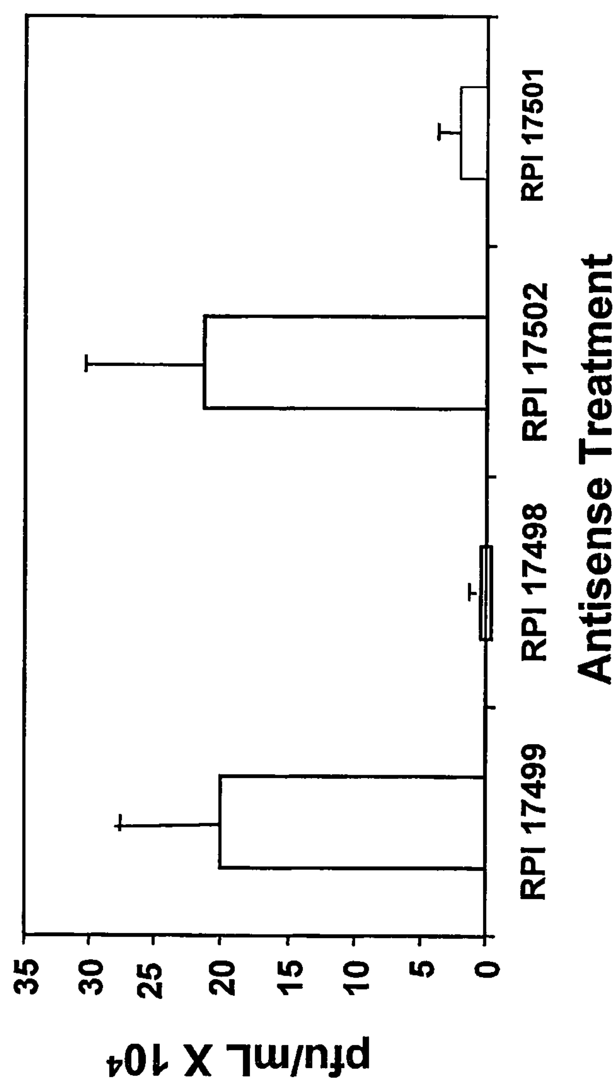
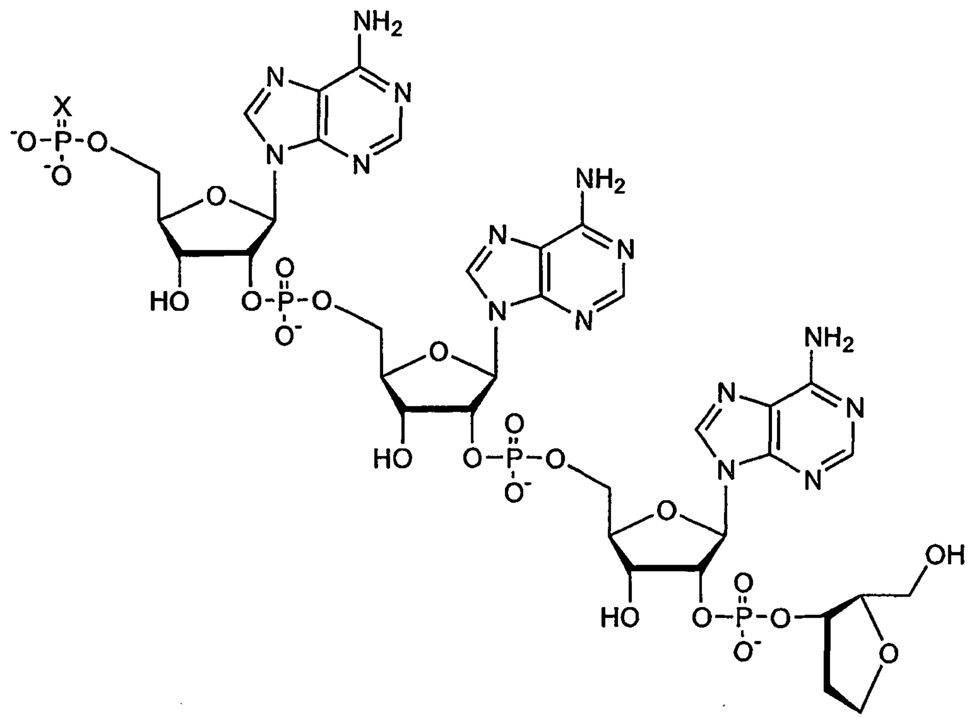


Figure 34: Inhibition of HCV-Poliovirus Replication by Antisense

Figure 35: Modified 2-5A Compound



I: X = O
II: X = S

Figure 36A: Ribozyme activity and enhanced antiviral effect

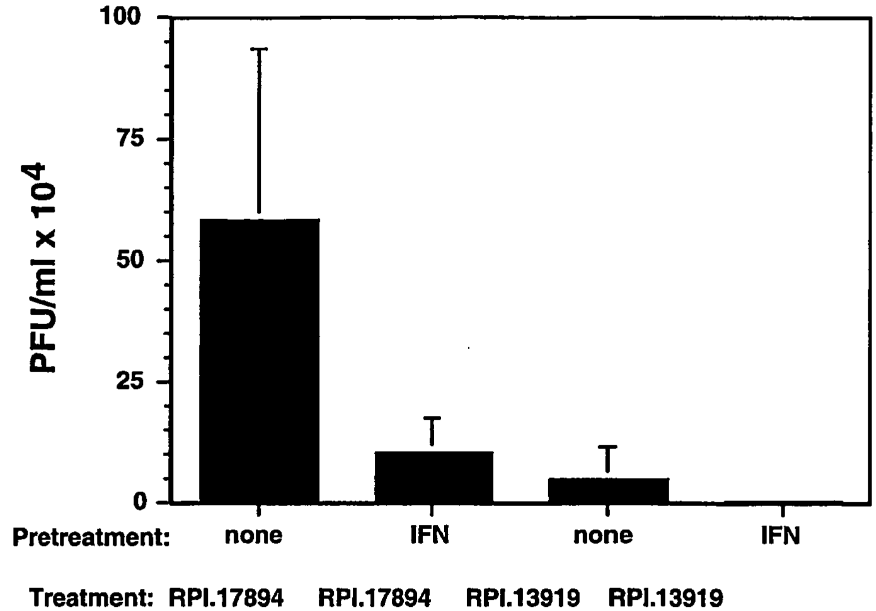


Figure 36B: Ribozyme activity and enhanced antiviral effect

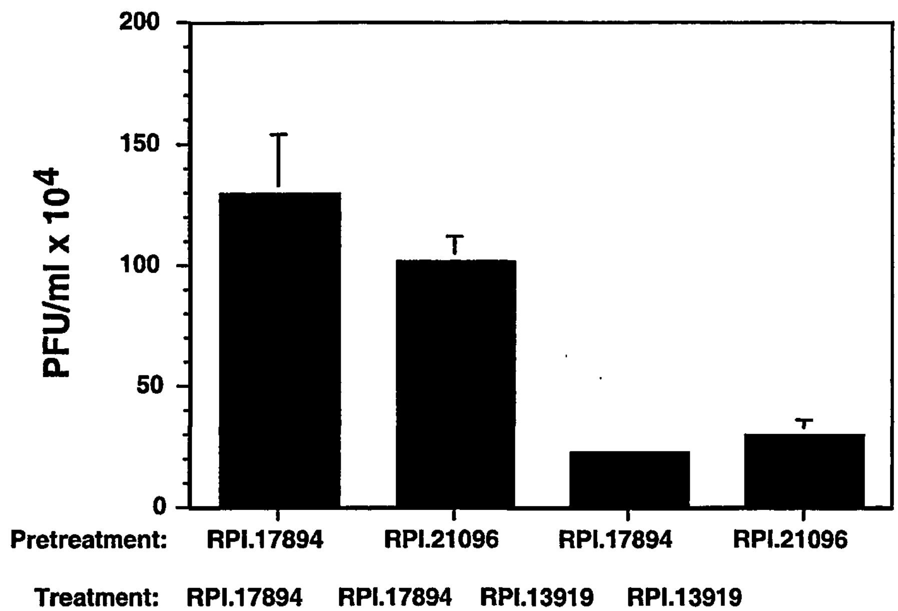


Figure 37: Inhibition of viral replication with anti-HCV ribozyme or 2-5A treatment

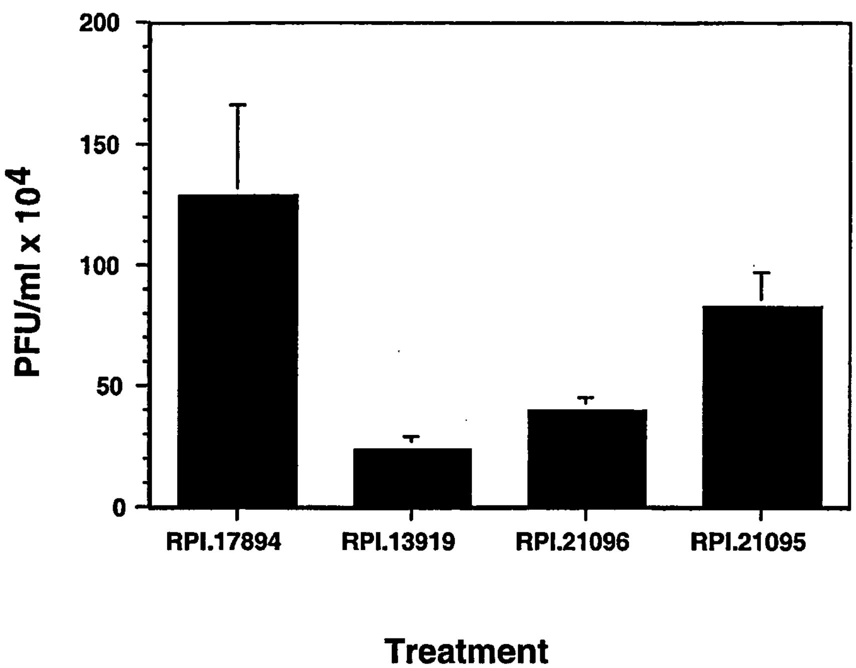
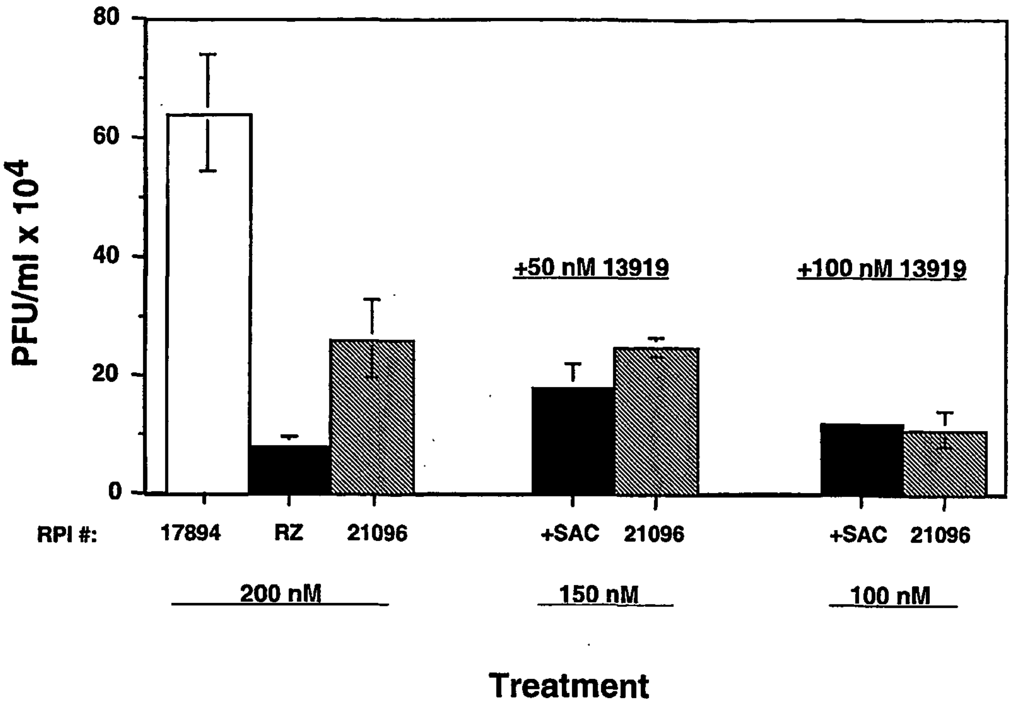


Figure 38: Anti-HCV ribozyme in combination with 2-5A treatment



This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ____________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.